Phenotypic and genotypic diversity of *Streptococcus pneumoniae* strains in Tanzania and the United Kingdom

By

Marcus Leung

A thesis submitted to University College London in fulfillment of the requirement for the degree of Doctor of Philosophy

Centre for Clinical Microbiology

Department of Infection

University College London

Royal Free Campus

Rowland Hill Street

London NW3 2PF

United Kingdom

Declaration

I, **Marcus Leung**, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in my thesis. In addition, figures and full-text publications throughout the thesis, where necessary, have been approved for reprint in this thesis via RightsLink.

Signed

Abstract

Streptococcus pneumoniae is a prevalent etiological agent of diseases and a common colonizer. During pneumococcal colonization, adaptation and evolution of the organism is thought to occur by horizontal gene transfer. Heterogeneity in distribution of genes between strains contributes to the global supragenome pool greater than the genome of any single strain. Co-colonization of multiple strains, therefore, enables an organism to expand its supragenome and hence its adaptive potential.

In this study, the development and design of a single primer-pair PCR- and sequence-based serotyping (sequetyping) method for surveillance of prevalent serotypes is presented. Serotypes including those covered in the latest conjugate vaccines were identified at least to the serogroup level. This method thus shows promise in its application in routine determination of prevalent serotypes.

The sequetyping method was applied to perform a systematic analysis of characterizing pneumococci isolated from healthy Tanzanian children in an attempt to determine phenotypic and genotypic diversities (multilocus sequence typing) within colonization events. Through this work evidence is presented to support that previous studies analyzing serotype diversity alone is an underestimate of true strain diversity during co-colonization events.

Determination of competence peptide allelic variants revealed significant differences in proportions of pherotypes in two geographical locations. The majority of co-

colonization events showed multiple pherotypes, indicating that co-colonization events may provide an optimal condition for genetic exchanges.

Intra-serotype variations in the capsular gene *cpsB* were observed and the difference of this variation between serotypes. Nucleotide differences were mapped to specific domains of the CpsB protein, which may affect the enzymatic activity of this regulatory protein.

Sequencing *comC* from presumptive strains of *S. pseudopneumoniae* revealed that *comC6.1* is a common pherotype of this organism. It is proposed that this gene may be a potential target for differentiation between *S. pseudopneumoniae* from *S. pneumoniae* and other streptococcal species.

Acknowledgements

I wish to acknowledge the support and guidance of my supervisors Dr. Bambos Charalambous [1] and Professor Stephen Gillespie [2]. Their invaluable insight and patience have truly made my Ph.D. experience rewarding, enjoyable, and worthwhile.

I would like to extend my thanks to Dr. Timothy McHugh for his advice and kindness through the years. I am grateful for the support and friendship of staff and students of the Centre for Clinical Microbiology Department of the Royal Free Medical School.

In addition I would like to thank the clinical scientists and laboratory personnel of the Microbiology Department of the Royal Free Hospital for their advice, courtesy, and friendship.

I am grateful for the help of the following individuals for their contribution to works included in this thesis: Ms. Kathrin Freystätter[1] and Dr. Kevin Bryson [3] for their design of the PrimerFinder algorithm, Dr. Bruno Pichon[4], Dr. Jutta Loeffler[5], Dr. Derren Ready[6], Dr. Anna Tymon[6] Ms. Rajita Dasai[7], Dr. Clare Ling[8], and Ms. Sarah Taylor[9], Dr. Ndekya Oriyo[10,] Dr. Sarah Batt[1], and the laboratory personnel of the Tanganyika Plantation Company Limited (TPC) Hospital and the Kilimanjaro Christian Medical College (KCMC) for providing bacterial strains. In particular, the

[1] Centre for Clinical Microbiology, University College London Medical School
[2] School of Medicine, University of St. Andrews
[3] Bioinformatics Group, Department of Computer Science, University College London
[4] Microbiology Services Division, Health Protection Agency, Colindale, UK
[5] Centre for Amyloidosis and Acute Phase Proteins, University College Medical School
[6] Microbial Diseases, Eastman Dental Hospital
[7] Microbiology Department, Royal Free Hampstead NHS Trust
[8] Microbiology Services, Health Protection Agency, Barts and the London NHS Trust
[9] Respiratory Medicine, University College London Medical School
[10] Kilimanjaro Christian Medical College, Moshi, Tanzania

collection, storage, and preliminary phenotypic characterization of strains from Tanzania were performed and supervised by Dr. Ndekya Oriyo, Dr. Sarah Batt and laboratory scientists of local hospitals.

I am indebted to the endless support and encouragement from my relatives, and friends from London, Vancouver, and Hong Kong, as they have provided me inspiration and motivation to continue and complete this thesis.

I would like to dedicate this thesis to my parents and sister, who have given endless support even when we were over 7,000 kilometers apart. Your love and understanding even at stressful and difficult times will forever be remembered. I love you all for this.

List of Abbreviations and Acronyms

AOM	Acute Otitis Media
bp	Base pair
CBD	Choline-Binding Domain
CBP	Choline-Binding Protein
CBR	Choline-Binding Region
CC	Clonal Complex
CGH	Comparative Genome Hybridization
CLSI	Clinical and Laboratory Standards Institute
CPS	Capsular Polysaccharide
CSF	Cerebrospinal Fluid
DNA	Deoxyribonucleic acid
ECM	Extracellular Matrix
GBS	Group B Streptococcus(cci)
HGT	Horizontal Gene Transfer
HIV	Human Immunodeficiency Virus
HPLC	High-Performance Liquid Chromatography
IPD	Invasive Pneumococcal Disease
kbp	Kilo Base pairs (1,000 bp)
KCMC	Kilimanjaro Christian Medical Centre
LD_{50}	Lethal Dose (50%)
LTA	Lipoteichoic Acid
MLST	Multilocus Sequence Typing
mRNA	Messenger Ribonucleic Acid
NMR	Nuclear Magnetic Resonance
NVT	Non-Vaccine (Sero)Type
OM	Otitis Media
OPA	Opsonophagocytic Assay
PBP	Penicillin-Binding Protein
Pce	Phosphorylcholine Esterase
PCho	Phosphorylcholine
PCR	Polymerase Chain Reaction
PCV7	Heptavalent Pneumococcal Conjugate Vaccine
PCV13	13-Valent Pneumococcal Conjugate Vaccine
PFGE	Pulsed-Field Gel Electrophoresis
PMEN	Pneumococcal Molecular Epidemiology Network
PspA	Pneumococcal Surface Protein A
RNA	Ribonucleic Acid
rRNA	Ribosomal Ribonucleic Acid
SNP	Single Nucleotide Polymorphism
ST	Sequence Type
TA	Teichoic Acid
TBE	Tris-Boric EDTA Buffer
TPC	Tanganyika Planting Company
tRNA	Transfer Ribonucleic Acid
URT	Upper Respiratory Tract
VT	Vaccine (Sero)Type

Table of Contents

List of Tables

List of Figures

Relevant Publications

CHAPTER THREE
Leung, M.H., K. Bryson, K. Freystatter, B. Pichon, G. Edwards, B.M. Charalambous, and S.H. Gillespie. Sequetyping: Serotyping *Streptococcus pneumoniae* by a single PCR-sequencing strategy. 2012. *J. Clin. Micro.* In press

CHAPTER FOUR
Leung, M.H. N.M. Oriyo, S.H. Gillespie, and B.M. Charalambous. The adaptive potential during nasopharyngeal colonization of *Streptococcus pneumoniae*. 2011 *Infect. Genet. Evol.* **11**: 1989-1995

CHAPTER FIVE
Leung, M.H., C.L. Ling, H. Ciesielczuk, J. Lockwood, S. Thurston, B.M. Charalambous, and S.H. Gillespie. *Streptococcus pseudopneumoniae* identification by pherotype: a method to assist the understanding of a potentially emerging or overlooked pathogen. 2012. *J. Clin. Micro.* **50**: 1684-1690

CHAPTER SIX
Leung, M.H., Schiebler, M., Batt, S., Crump, J.A., Lake, D., Gillespie, S.H., and Charalambous, B.M. *An analysis of United Kingdom and Tanzania pherotypes found in carriage and clinical strains of* Streptococcus pneumoniae. Submitted to PLoS Pathogens

Other Publication

Charalambous B.M., and M. H. Leung. *Pneumococcal sepsis and nasopharyngeal carriage*. 2012. *Curr. Opin. Pulm. Med.* **18**: 222-227.

CHAPTER ONE: General Introduction

1.1 The Organism

1.1.1 History

Streptococcus pneumoniae was first isolated in 1880 when George Miller Sternberg inoculated his own saliva into a rabbit and isolated diplococci in its blood, followed by Louis Pasteur's discovery of the same organism few months later in blood after inoculating a rabbit with saliva from a child who died of rabies (Watson *et al.* 1993). Pasteur named the organism *Microbe septicemique du salive* and Sternberg termed the organism *Micrococcus pasteuri.* The organism was later known as *Diplococcus pneumoniae* as isolates from pneumonia appeared as pairs. Its current scientific name, *Streptococcus pneumoniae*, was adopted based on the chain-like growth of pneumococci in liquid media (Diebel & Seeley 1974). The history of *S. pneumoniae* is reviewed in numerous publications (Austrian 1981; Watson *et al.* 1993; Watson & Musher 1999; López 2006).

Some of the most profound biological principles and inventions were tested and proven using the pneumococcus, such as invention of Gram staining, the concept of opsonisation, the discovery of polysaccharides as possible antigenic substances, and the discovery of penicillin-binding proteins[11] (Section 1.3.2). The greatest scientific discovery of the 20th century where the pneumococcus was tested was perhaps the transforming principle, which proved DNA to be the hereditary material (see below).

1.1.2 The Transformation Principle

Frederick Griffith identified a live rough (R) and smooth (S) form of pneumococcus, where only the latter expressed the virulent polysaccharide capsule (Griffith 1928). When a live, avirulent R strain was injected subcutaneously with a heat-killed pathogenic S strain into mice, he found that the R strain had "transformed" to express the same capsule type as the heat-killed strain. Oswald Avery, Colin MacLeod, and Maclyn McCarty (Avery, Macleod & McCarty 1944) continued this work and

[11] PBPs are also expressed in species of *Neisseria* spp., *Escherichia* spp., *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Haemophilus influenzae*, etc. (Georgopapadakou 1993).

provided evidence concluding that the active transforming material was DNA and not protein. This landmark discovery paved way for a plethora a studies understanding DNA and genes, and consequently to the rapid advances in molecular biology.

1.1.3 Bacteriology and Microbiology

1.1.3.1 Growth Conditions and Characteristics of Streptococci

Members of the *Streptococcus* genus belong to the Firmicutes phylum, Lactobacillales order, and Streptococcaceae family. They are Gram-positive facultative anaerobe mainly seen in pairs and chains. They are non-spore forming and non-motile. They generate lactic acid from glucose fermentation. They are catalase-negative, and the addition of exogenous catalase in growth media is required to neutralise the toxic effects of hydrogen peroxide synthesized by the organism. Growth is optimal on agar supplemented with blood in at 35-37°C when supplied with 5% CO_2. They undergo either α-, β-, or γ- haemolysis[12]. Members of the *Streptococcus* genus, classified by their groups, are indicated in Table 1.1.

[12] α-haemolysis is the reduction of haemoglobin to methemoglobin by hydrogen peroxide, β-haemolysis is the breakdown of red blood cells, γ-haemolysis is non-haemolytic. The Mitis group of Streptococci, including *S. pneumoniae*, is α-haemolytic.

Table 1.1 Members of the *Streptococcus* genus, divided by genus subgroups

Group	Species
Mitis	***Streptococcus pneumoniae***
	Streptococcus mitis
	Streptococcus oralis
	Streptococcus pseudopneumoniae
	Streptococcus infantis
	Streptococcus cristatus
	Streptococcus gordonii
	Streptococcus sangiunis
Pyogenic	*Streptococcus pyogenes*
	Streptococcus agalactiae
	Streptococcus dysgalactiae
	Streptococcus iniae
	Streptococcus uberis
	Streptococcus equi
Anginosus	*Streptococcus anginosus*
	Streptococcus intermedius
	Streptococcus constellatus
Salivarius	*Streptococcus salivarius*
	Streptococcus thermophilus
	Streptococcus vestibularius
Bovis	*Streptococcus bovis*
	Streptococcus gallolyticus
	Streptococcus infantarius
	Streptococcus equinus
Mutans	*Streptococcus mutans*
	Streptococcus sobrinus
	Streptococcus downeii
	Streptococcus rattus
	Streptococcus cricetus

1.1.3.2 Phenotypic Pneumococcal Identification

S. pneumoniae is in the Mitis group of streptococci with a cell diameter of 0.5 – 1.25 μm, and its colonies have a glistening and mucoid appearance on blood agar plates (Kawamura *et al.* 1995). Depression at the centre of colonies (draughtsman's shape) can be seen due to autolysis during stationary phase. Blood agar plates containing gentamicin may be used for selection of pneumococci from naso-/oropharyngeal samples where other organisms may be in abundance (Sondag *et al.* 1977; Dilworth *et*

al. 1975; Dudley *et al.* 2001). Most pneumococci express one of 93 structurally variable polysaccharide capsules (Section 1.1.4.1), which are differentiated by serological and DNA-based methods (Chapter 3).

The pneumococcus is differentiated from other related (viridans) streptococci by α-haemolysis, colony morphology, optochin (ethylhydrocuprein hydrochloride) sensitivity, bile (deoxycholic acid) solubility, and agglutination with capsular-specific antibodies. Pneumococci are generally optochin-sensitive and bile-soluble. However, optochin-resistant pneumococci (Pikis *et al.* 2001; Ing *et al.* 2012), optochin-sensitive viridans streptococci (Borek *et al.* 1997; Balsalobre *et al.* 2006) and bile-soluble viridans streptococci (Messmer, Black & Facklam 1995) complicate conventional classification. This is further aggravated with the recent identification of *Streptococcus pseudopneumoniae* (Chapter 7), a species that is bile-insoluble and acapsulate, but is optochin-resistant in presence of increased (5%) CO_2 and optochin-sensitive in ambient air (Arbique *et al.* 2004). Biochemical tests such as the Rapid ID32 Strep System and the API Rapid Strep Test (bioMerieux) have been used for species differentiation, however results are inconclusive and additional tests are often required to confirm species identity (Freney *et al.* 1992; Kikuchi *et al.* 1995; Arbique *et al.* 2004).

1.1.3.3 Genetic Pneumococcal Differentiation from Viridans Streptococci

S. pneumoniae shares 99% 16S rRNA sequence similarity with *S. mitis* and *S. pseudopneumoniae* (Kawamura *et al.* 1995; Arbique *et al.* 2004), However, recent reports suggest that internal species-specific signature sequences of the 16S rRNA gene can distinguish these species (El Aila *et al.* 2010). Autolysin and pneumolysin were thought to be exclusive pneumococcal virulence factors. However, viridans streptococci with *lytA* and *ply* homologues have been detected (Whatmore *et al.* 2000; Neeleman *et al.* 2004; Jefferies *et al.* 2007a; Johnston *et al.* 2010). DNA-based differentiation methods include sequence analysis of housekeeping genes (Kawamura *et al.* 1999; Hanage *et al.* 2005; Kilian *et al.* 2008). Other genetic targets for species differentiation include Spn9802, Spn9828, *cpsA*, *recA*, and *psaA* (Suzuki *et al.* 2005; El Aila *et al.* 2010; Park *et al.* 2010; Sistek *et al.* 2011). Correct identification of streptococcal species is important, as misidentification of pneumococci based on optochin sensitivity, bile solubility, and non-specific markers potentially leads to an

overestimation of prevalence in pneumococcal antibiotic-resistance and hence incorrect use of antibiotics and obscures the pathogenic potentials of viridans streptococci (Wester *et al.* 2002; Neeleman *et al.* 2004).

1.1.4 Cellular Components

The pneumococcus expresses a polysaccharide capsule on its outermost layer, followed by a peptidoglycan cell wall containing teichoic and lipoteichoic acids (TA and LTAs). Pneumococcal surface proteins are bound to the cell surface either via the choline residues on TAs and LTAs, peptidoglycan chains, or cell membrane (Fig. 1.1). Other major proteins may be cytoplasmic. Selected cellular structures are described below.

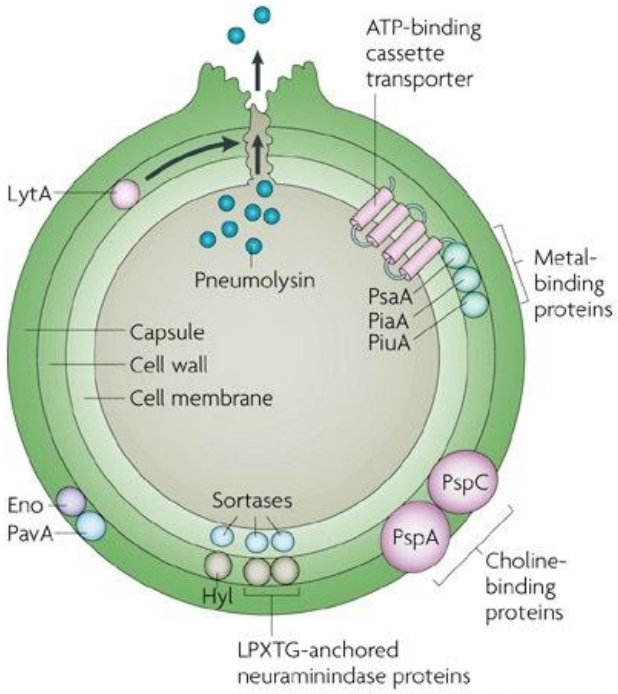


Fig. 1.1. Pneumococcal virulence and colonization factors. Representatives of surface proteins indicated based on their method of association to cell wall (metal-binding, LPxTG-motif cell wall associated, choline-binding proteins). Additional proteins from each of these types are described below, as well as the cytoplasmic pneumolysin, and cell-surface carbohydrate moieties (LTA/TA, capsule). Figure from Kadioglu *et al.* (Kadioglu *et al.* 2008) with permission granted from Nature Publishing Group via RightsLink Copyright Clearance Center.

1.1.4.1 Polysaccharide Capsule

The capsule is a high molecular mass complex of oligosaccharide repeating units covalently linked to the peptidoglycan layer (Sørensen *et al.* 1990). The thickness of the capsule ranges from 200-400 nm. The capsule has been considered as the *sine qua non* of virulence, and strains with defects in capsular production are either avirulent or attenuated in infection and colonization models (Watson & Musher 1990; Bender & Yother 2001; Hardy *et al.* 2001; Magee & Yother 2001; Morona *et al.* 2004; Nelson *et al.* 2007b). The polysaccharide allows the organism to evade complement-mediated opsonophagocytosis and removal from the mucosal surface (Kim *et al.* 1999; Nelson *et al.* 2007b; Hyams *et al.* 2010), and strains with higher levels of capsule expression are more resistant to phagocytosis (Weinberger *et al.* 2009). Acapsulate strains rarely cause diseases except in cases of conjunctivitis (Section 1.2.1). The presence of the capsule may also affect differential expression of immunomodulatory genes by the host epithelium during attachment, thereby affecting the host response (Bootsma, Egmont-Petersen & Hermans 2007). The regulation of capsule production has been shown to be crucial for pneumococcal survival within the host; full encapsulation appears to be beneficial for systemic virulence, while reduced expression of capsular polysaccharide aids adherence and biofilm formation (Magee & Yother 2001; Hammerschmidt *et al.* 2005; Moscoso, García & López 2006).

The pneumococcus undergoes phase variation between opaque and transparent forms varying in opacity and colony morphology, with a frequency ranging from 10^{-3} to 10^{-6} depending on the strain (Weiser *et al.* 1994). Opaque variants produce up to 5.6 times more CPS than transparent variants, while transparent variants were reported to produce up to 3.8 times more teichoic acid (Kim & Weiser 1998). Transparent variants are more efficient in stable colonization and clonal expansion in the nasopharynx in infant rats, while opaque variants are frequently associated with systemic diseases.

1.1.4.1.1 *Capsule Biochemistry*

Ninety-three serotypes have thus far been identified (Calix & Nahm 2010). Serotypes can be categorized by two systems; the American system ranks serotypes by order of discovery, while the more common Danish system groups cross-reactive serotypes into serogroups. Under the Danish system, forty-six serotypes/groups are numbered

from 1 to 48 (numbers 26 and 30 not used) based on their structural properties and immunogenicity. Serotypes can be differentiated and identified by type-specific anti-capsular rabbit antibodies (Chapter 3). Cross-reactive serotypes are categorized into a serogroup and given a letter (e.g. serotypes 6A, 6B, 6C, 6D within serogroup 6). The primary structures of the capsule vary between serotypes by the types and number of monosaccharides present, the monosaccharide sequences and ring sizes, glycosidic linkage configurations, and non-sugar elements (Kamerling 2000). Most serotypes have negatively charged capsules due to the presence of acidic sugars, pyruvate and/or phosphate groups. Serotypes 7F, 7A, 14, 33F 37 have no net charge. Serotype 1 has a zwitterionic charge containing positive and negative charges on the backbone substituents.

The chemical structure of each capsular serotype may be associated with their ability to colonize the host; Weinberger *et al.* (Weinberger *et al.* 2009) demonstrated that serotypes with fewer number of carbon repeats (e.g. 6A, 6B, 19F, 23F) produce more capsules, are more resistant to neutrophil-mediated phagocytosis, and are more prevalent in carriage. The authors also predict that analysing capsule structures can also assist in predicting which serotypes will be involved in serotype replacement (Weinberger *et al.* 2009).

Capsule enzymes are termed either by the Wzy system (based on homologues of *E. coli*), or by Cps (capsular polysaccharide synthesis) followed by a letter based on the order of the enzyme/gene from 5' to 3'. Table 1.2 shows the proteins of the *cps* operon and the corresponding names in the Wzy nomenclature system based on the genetic structure of serotype 19F (Muñoz, Mollerach & López 1997; Bentley *et al.* 2006).

Table 1.2. Capsular polysaccharide proteins of serotype 19F based on Wzy and Cps nomenclatures

Biological Role	Wzy Nomenclature	Cps Nomenclature
Regulatory	Wzg	CpsA
	Wzh	CpsB
	Wzd	CpsC
	Wze	CpsD
Initial transferase	WchA	CpsE
Glycosyl transferase	WchO	CpsF
	WchP	CpsG
	WchQ	CpsH
Polymerase	Wzy	CpsI
Flippase	Wzx	CpsJ
Sugar biosynthesis	MnaA	CpsK
	RmlA	CpsL
	RmlC	CpsM
	RmlB	CpsN
	RmlD	CpsO

1.1.4.1.2 *Capsule Assembly*

All serotypes except 3 and 37 follow the Wzy-dependent (Block-Type) capsular synthesis pathway (Fig. 1.2), a fashion similar to capsule production in other streptococci and staphylococci, as well as lipopolysaccharide O-antigen synthesis in Group 1 *Escherichia coli* (Whitfield & Paiment 2003; Yother 2011). For these serotypes, biosynthesis of nucleotide-sugar is initiated in the cytoplasm and is transferred to the undecaprenyl phosphate lipid acceptor, catalyzed by initial glycosyltransferases (Kolkman, van der Zeijst & Nuijten 1998; Yother 2004; Cartee *et al.* 2005). Extension of the repeating unit is catalyzed by additional membrane-associated glycosyltransferases (Aanensen *et al.* 2007). The repeat unit transporter, also known as the flippase (Wzx), then translocates the repeat unit from the cytoplasmic to the periplasm, followed by the polymerization (by Wzy polymerase) of the repeating units to form a lipid-linked polysaccharide capsule. Lastly, information regarding the release of the capsule from the lipid layer and attachment to cell wall remains elusive, but may involve CpsC (Wzd) and CpsD (Wze) (Morona, Morona & Paton 2006).

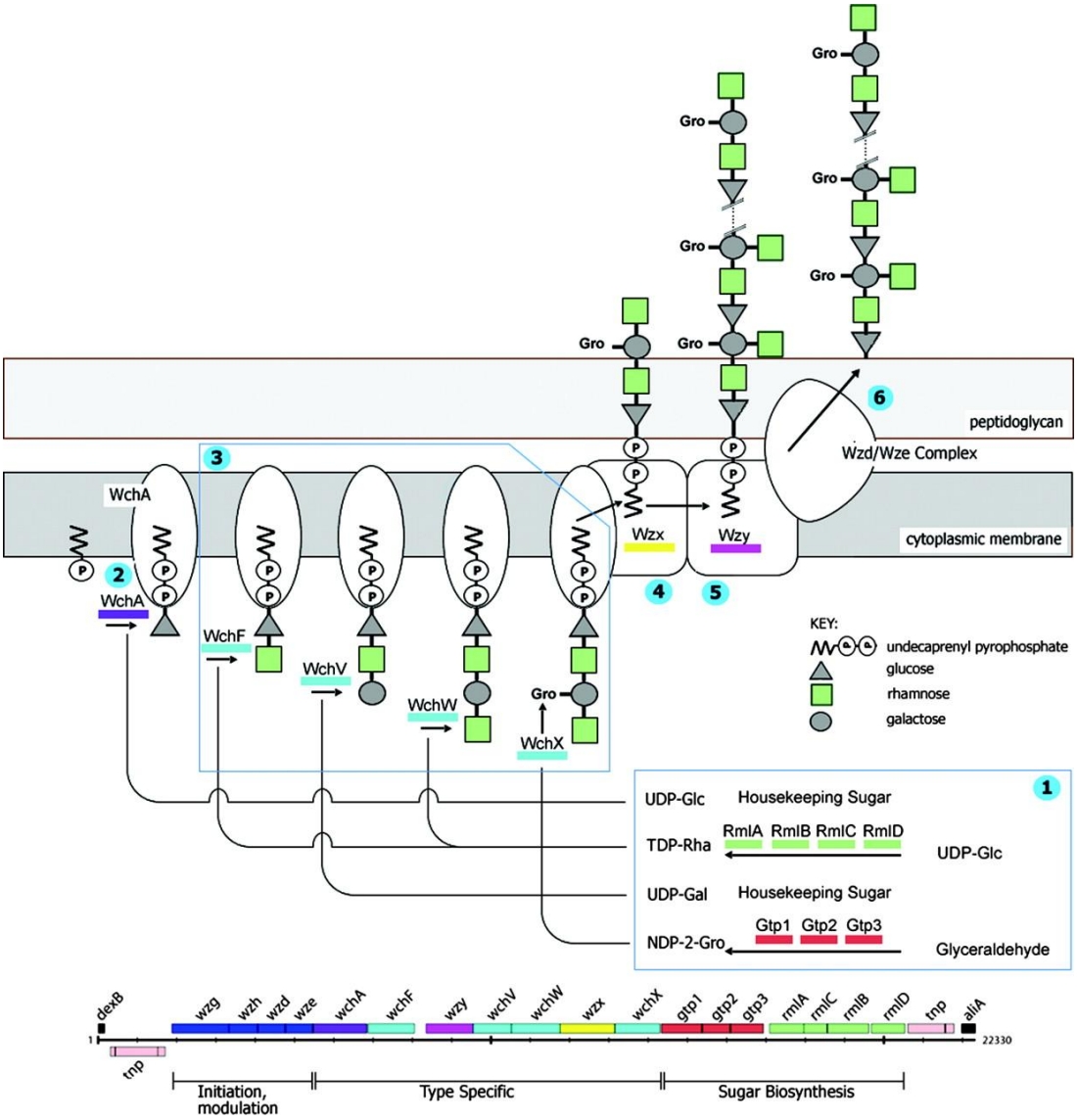


Fig. 1.2. Capsule assembly via Wzy-dependent mechanism. Diagram depicts assembly process for serotype 9A. Capsule assembly is as follows: 1) the biosynthesis of non-housekeeping sugars by proteins encoded by genes in the 3' end of the operon, 2) the transfer of first sugar of the repeat unit onto a lipid-linked anchor, mediated by the initial glycosyltransferase (WchA in serotype 9A), 3) the addition of subsequent saccharides to form a repeating unit mediated by additional glycosyltransferases, 4) the transfer of repeating unit from the cytoplasmic space to the periplasm, mediated by flippase Wzx, 5) the polymerization of repeating units by polymerase Wzy, 6) the attachment of capsule polymer to the peptidoglycan via Wzd/Wze complex. Figure from Aanensen *et al.* (Aanensen *et al.* 2007). Journal of Bacteriology authorizes the reuse of journal material including figure shown for academic purposes.

Serotypes 3 and 37 use the synthase-dependent synthesis pathway. For serotype 3, synthesis is initiated by linking glucose (Glc) and glucuronic acid (GlcUA) to form the capsule polymer by a membrane-associated glucosyltransferase (synthase) encoded by *cps3S* (Llull *et al.* 1999; Cartee *et al.* 2000) (Fig. 1.3). The synthesis of the serotype 37 glucose homopolymer capsule is mediated by a single β-glucosyltransferase encoded by the *tts* gene located outside the *cps* locus (Llull *et al.* 1999).

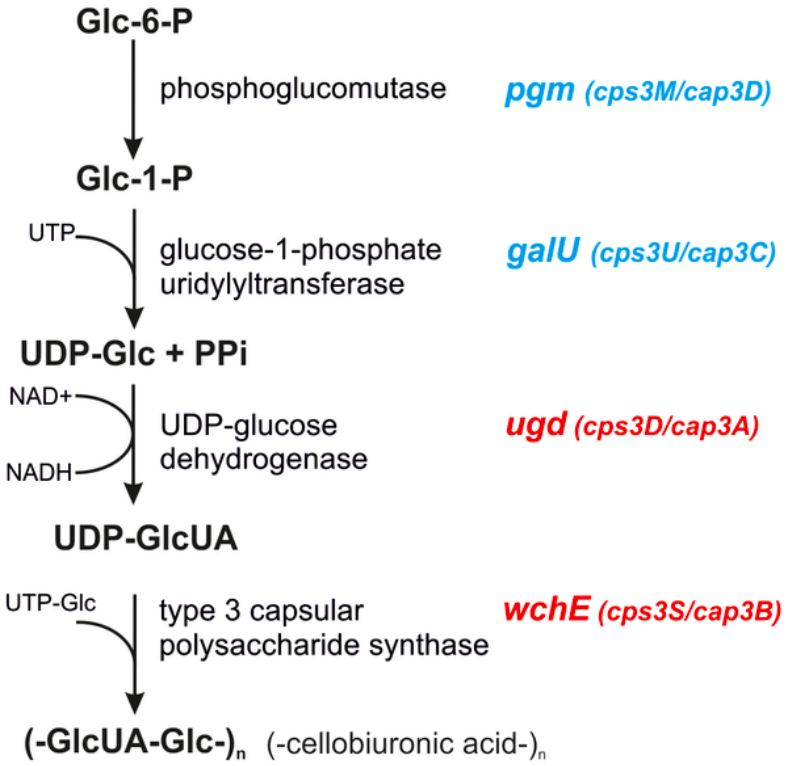


Fig. 1.3. Biochemical pathway of capsule assembly via synthase-dependent mechanism (serotype 3). Pathway illustrated for synthesis of serotype 3 capsule. Figure from Denapaite and Hakenbeck (Denapaite & Hakenbeck 2011). PLoS grants unrestricted use of contents including figure shown.

1.1.4.1.3 *Genetic Structure of cps Loci for Wzy-Dependent Pathway Serotypes*

Most genes required for capsule synthesis are within the capsule polysaccharide synthesis (*cps*) operon (Fig. 1.2) (Bentley *et al.* 2006). The operon is located in the same chromosomal region in different serotypes, and ranges from 10 kb (serotype 3) to 30 kb (serotype 38). In all serotypes, *cps* is flanked by *dexB* in 5' end and *aliA* in the 3' end, neither of which participates in capsule synthesis. The 5' end of the locus, encoding *cpsA*, *cpsB*, *cpsC*, and *cpsD*, are conserved in all serotypes and is involved in capsule production regulation and perhaps attachment of capsule to the cell wall.

The next gene is *cpsE* encoding the initial glycosyltransferase, catalyzing the transfer of glucose-1-phosphate to the lipid-linked undecaprenyl-phosphate moiety in most serotypes that contains glucose in their capsules. The central region of the locus encodes specific enzymes required for the transfer of sugar constituents of a repeating unit. Within this region, the flippase (*wzx*) and polymerase (*wzy*) genes mediate the export and polymerization of the repeating units at the extracellular space. The diversity of gene sequences and gene content in this region is a reflection the enzymes' specificities for substrates that constitute the capsule. Genes at the 3' end of the locus encodes for proteins required for non-housekeeping capsule sugar biosynthesis. Housekeeping sugars present in the capsule are present elsewhere in the chromosome (Jiang, Wang & Reeves 2001; Bentley *et al.* 2006).

The locus contains a functional conserved promoter immediately upstream of *cpsA*. The operon is transcribed as a single unit, with the transcriptional start site eight nucleotides after the -10 consensus region of the promoter (Muñoz, Mollerach & López 1997; Llull *et al.* 1998). Transposase sequences are present between *cps* genes and *dexB*/*aliA* that may be involved with serotype switching (Section 1.4.4). In serotypes 25A, 25F, and 38, the rearrangement of the 5' genes is associated with the presence of transposase genes between *cpsCD* and *cpsAB* (Muñoz, Mollerach & López 1997; Mavroidi *et al.* 2007). Cryptic synthesis genes for a particular monosaccharide may be present even though the monosaccharide is not expressed on the capsule; ORFs coding for dTDP-rhamnose synthesis have been found in serotype 1, a serotype that does not have rhamnose in its capsule (Muñoz, Mollerach & López 1997).

Genes with genetic similarity to the pneumococcal capsulation loci have also been detected in other capsulated streptococci such as *S. agalactiae*, *S. mitis*, *S. oralis*, *S. suis*, and *S. thermophilus* (Smith *et al.* 1999; Cieslewicz *et al.* 2001; Mavroidi *et al.* 2007; Rukke, Hegna & Petersen 2012).

1.1.4.1.4 *Genetic Structure of cps Loci for Synthase-Dependent Pathway Serotypes*

Dillard *et al.* (Dillard & Yother 1994; Dillard, Vandersea & Yother 1995) reported the serotype-specific genes *cps3D*, *cps3S*, and *cps3U* (or *cap3A*, *cap3B*, and *cap3C*, respectively, according to Arrucubieta *et al.* (Arrecubieta, López & García 1994;

Arrecubieta, Garcia & López 1995) (Fig. 1.4). The *cap3ABC* genes are not to be confused with the regulatory genes *cpsABC* of other serotypes as *cap3A*, *cap3B* and *cap3C* are serotype-specific genes required for capsule polymerization. A promoter is located upstream of *cpsD*. Further upstream are sequences similar to that of regulatory genes of other serotypes, however deletions are present in this region, and these pseudogenes are not transcribed in serotype 3 (Arrecubieta, Garcia & López 1995). The first 3 genes appear to be transcribed as a single unit, but only *cpsD* and *cpsS* in this operon has been shown to be required for encapsulation (Arrecubieta, Garcia & López 1995). Also, *cpsD*-knockout strains were able to express serotype 3 capsules if UDP-Glc and UDP-GlcA were supplied *in vitro* (Dillard, Vandersea & Yother 1995).

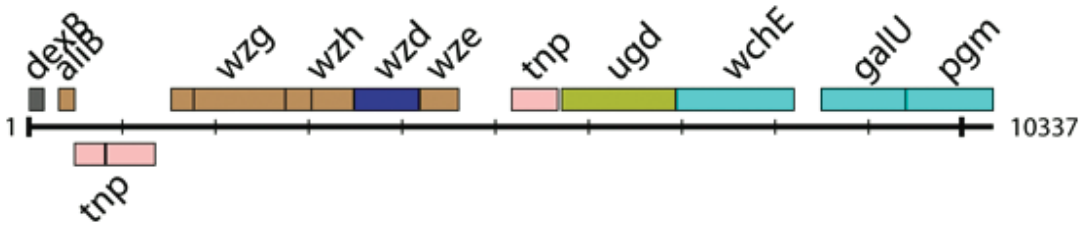


Fig. 1.4 Genetic organization of *cps* in serotype 3. The genes *ugd* (glucose dehydrogenase) *wchE* (glycosyltransferase), *galU* (pyrophosphorylase), and *pgm* (phosphoglucomutase) are *cpsD*, *cpsS*, *cpsU*, and *cpsM* respectively by the *cps* nomenclature. All upstream genes except for *wzd* contain mutations (i.e. pseudogene) and are not expressed. Transposase genes indicated as *tnp*. Figure from Bentley *et al.* (Bentley *et al.* 2006). PLoS grants unrestricted use of contents including figure above for academic purposes.

The *cps* of serotype 37 is not required for capsule production. Instead, capsule synthesis in this serotype requires a single *tts* gene located outside the *cps* locus (Llull *et al.* 1999). Between *dexB* and *aliA* in serotype 37 is a cryptic *cps33f* locus, containing multiple mutations inactivating *cps37B*, *cps37E*, *cps37N*, and *cps37O*.

1.1.4.2 Cell Wall

The pneumococcal cell wall, found below the capsule polysaccharide, is a complex cellular layer containing a dense network of peptidoglycan (murein), teichoic acid (TA), and lipoteichoic acid (LTA) (Fig. 1.5). TA and LTA are decorated with phosphrylocholine (PCho) residues acting as anchoring sites for choline-binding proteins (CBP) (see below).

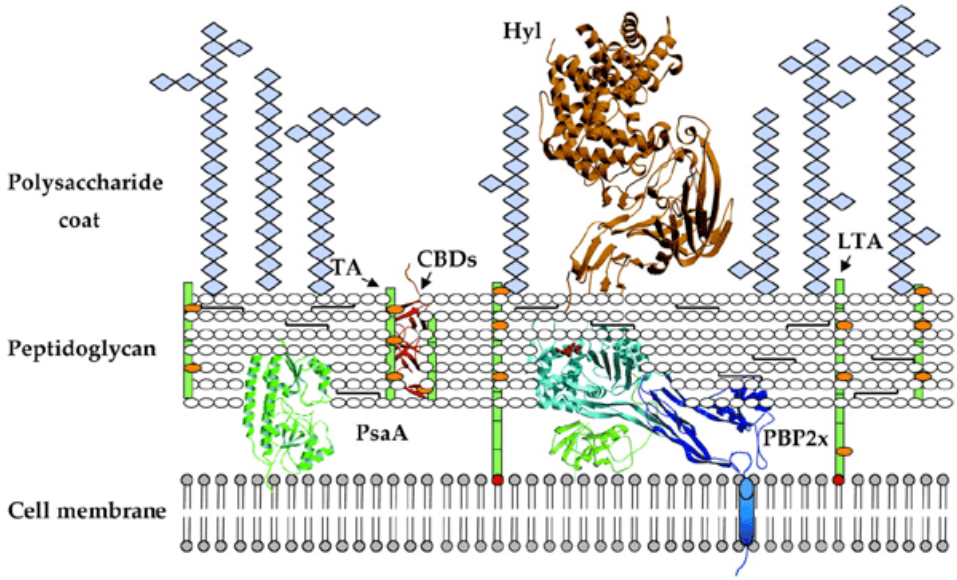


Fig. 1.5. Schematic diagram of pneumococcal cell wall components. Figure from Di Guilmi and Dessen (Di Guilmi & Dessen 2002) with permission granted from Nature Publishing Group via RightsLink Copyright Clearance Center.

1.1.4.3 Peptidoglycan

Peptidoglycan, sometimes referred to murein, is a multi-layered network of glycan strands interconnected by short peptide chains. A typical glycan strand consists of alternating *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) residues. Up to 80% of the Glc*N*Ac may be deacetylated by a deacetylase (PdgA), forming glucosamine. PdgA has been shown to play a role in resistance to lysozyme-mediated degradation of the cell wall by the host (Vollmer & Tomasz 2000).

Peptidoglycan in *S. pneumoniae* has cross-linked short peptide chains connected via the Mur*N*Ac residue and the terminal alanine residues of the stem peptide. There are 18 different stem muropeptide chains found highly conserved in pneumococcus of different serotypes, geographical origin, and isolation date (Severin & Tomasz 1996). Differences in proportions of these stem peptide compositions between strains may be due to concentrations of various amino acids in growth medium, or amino acid substitutions in PBP active sites associated with nonsusceptibility to penicillin (Section 1.3.2).

1.1.4.4 Teichoic Acid (TA) and Lipoteichoic Acid (LTA)

TA, also known as C-polysaccharide, is covalently linked to *N*-acetylmuramyl residues of peptidoglycan via a phosphodiester bond. LTA, also known as F-antigen, is uniformly distributed around the cell wall, anchored to the plasma membrane by the fatty acid component of the molecule (Fischer 2000). TA and LTA have identical repeating units, are complex in chemical structure, and contain choline acting as anchor for a wide array of pneumococcal surface proteins (see below). TA and LTA are common pneumococcal antigens, with positive reaction to anti-C-polysaccharide antiserum in 90 different capsular serotypes (Sørensen & Henrichsen 1987; Henrichsen 1995). A schematic representation of LTA and TA is provided in Fig. 1.6.

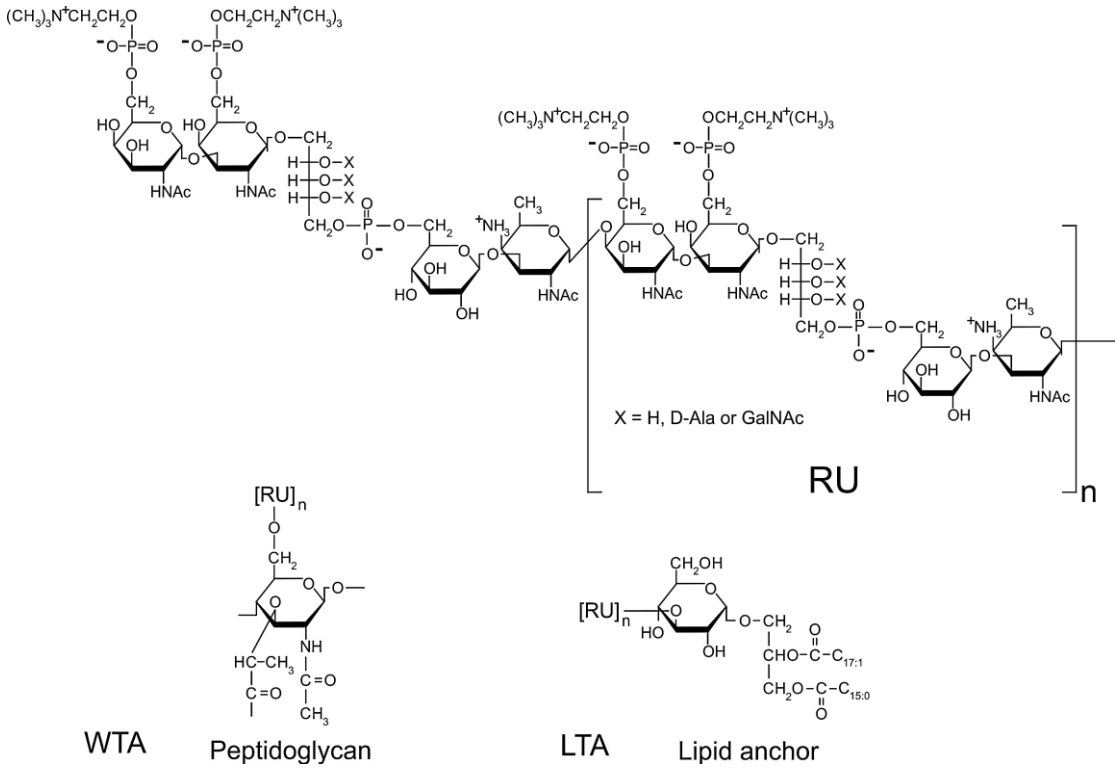


Fig 1.6. Structures of TA and LTA. LTA and TA (indicated here as WTA for wall teichoic acid) have identical repeating units. The repeating units (RU) are identical in both TA and LTA. In TA, the RU is linked to the peptidoglycan via the Mur*N*Ac, while in LTA the RU is linked to a fatty acid component (lipid anchor). Figure from Denapaite *et al.* (Denapaite *et al.* 2012).

1.1.4.5 Phosphorylcholine

The pneumococcal cell wall contains phosphorylcholine (PCho) covalently linked to TAs and LTAs. PCho was first found to be associated with pneumococcal cell wall in 1967 (Tomasz 1967), and since then it had been discovered in other bacterial species including nasopharyngeal colonizers *S. mitis*, *S. oralis*, *Neisseria meningitidis*, and *Haemophilus influenzae* (Weiser *et al.* 1998a; Weiser *et al.* 1998b). PCho promotes colonization and invasion by anchoring an array of choline-binding proteins (CBPs), many of which are virulence factors and adhesins crucial for pathogenicity and attachment (Reviewed in Jedrzejas 2004). PCho also directly binds to the platelet-activating factor receptor (PAFr) of stimulated host epithelial cells, playing a role in pneumococcal transcytosis through host cells during disease progression (Cundell *et al.* 1995; Ring, Weiser & Tuomanen 1998).

Mutants lacking choline uptake mechanisms showed marked reduction in adherence to nasopharyngeal epithelial cells, as well as deficiencies in colonization and invasion in mice (Tomasz 1968; Kharat & Tomasz 2006). In addition, mutants growing in choline-free medium appear as long chains and are resistant to autolysis, bile-insoluble, and genetically incompetent.

1.1.5 Pneumococcal Proteins

S. pneumoniae expresses numerous proteins crucial to colonization and virulence. Some, if not all, of these proteins are crucial in the pathogenicity of the organism, and are hence considered as virulence factors. Such virulence factors are reviewed in several reports (Gillespie & Balakrishnan 2000; Jedrzejas 2001; Jedrzejas 2004; Kadioglu *et al.* 2008; Mitchell & Mitchell 2010). The proteins have been organized in this section based on localization with respect to the cell: 1) **cytoplasmic proteins**, 2) **choline-binding proteins (CBPs)** associate with TA/LTA phosphorylcholine residues via non-covalent interactions, 3) **peptidoglycan-associated proteins** anchor to peptidoglycan via its LPxTG motif, and 4) **lipoproteins**. In addition, some pneumococcal proteins are anchored to the cell surface via yet unknown mechanisms.

1.1.5.1 Cytoplasmic Protein - Pneumolysin

The 471-amino acid, 52.8-kDa thiol-activated toxin binds to the host cholesterol receptor and subsequently forms an oligomeric pore (Morgan *et al.* 1994; Rossjohn *et*

al. 1998). Pneumolysin has been shown to damage the blood-brain barrier disrupting the integrity of microendothelial cells (Zysk *et al.* 2001). Pneumolysin activates phospholipase A_2 in endothelial cells, releasing cytolytic free fatty acids and lysophosphatides to induce inflammation (Rubins *et al.* 1994). Pneumolysin-deficient strains are attenuated in animal models of infection (Berry & Paton 2000; Wellmer *et al.* 2002), but animal studies provide conflicting observations regarding pneumolysin's role in colonization (Rubins *et al.* 1998; Kadioglu *et al.* 2002; van Rossum, Lysenko & Weiser 2005). Expression of pneumolysin may promote bacterial clearance from the nasopharynx in mice (van Rossum, Lysenko & Weiser 2005), and pneumolysin-deficient strains had higher bacterial count and longer duration of carriage in the nasopharynx (Rubins *et al.* 1998). However, one other work showed that pneumolysin expression was associated with higher bacterial count in the nasopharynx of mice (Kadioglu *et al.* 2002), however these studies employed strains of different serotypes and genetic backgrounds.

At high concentrations (1 μg/mL), pneumolysin is capable of activating the classical complement pathway, depleting complement components during infection (Paton, Rowan-Kelly & Ferrante 1984; Mitchell *et al.* 1991; Alcantara, Preheim & Gentry-Nielsen 2001). In nano-molar concentrations, purified pneumolysin impedes the respiratory burst of polymorphonuclear leukocytes (PMNLs), reducing antimicrobial activity (Paton & Ferrante 1983). Additionally, pneumolysin significantly stimulates the production of cytokines and nitric oxide in monocytes (Houldsworth, Andrew & Mitchell 1994)) and macrophages (Braun *et al.* 1999) *in vitro*, respectively. Allelic variants with differing hemolytic activity have been detected, and those with lowered activity were associated with strains causing IPD (ST306 of serotype 1 and ST53 of serotype 8) (Lock *et al.* 1996; Kirkham *et al.* 2006; Jefferies *et al.* 2007b; Jefferies *et al.* 2010b; Harvey *et al.* 2011).

Initial data supported the hypothesis that the release of pneumolysin is dependent on autolysin (Berry *et al.* 1989). However, the release of pneumolysin with inactivated autolysin and *lytA*-deletion suggests that additional, autolysin-independent mechanisms potentiate pneumolysin release (Balachandran *et al.* 2001).

1.1.5.2 Choline-Binding Proteins

1.1.5.2.1 *Autolysins (LytA, LytB, LytC)*

The autolysins form a group of murein hydrolases, cleaving covalent bonds of the peptidoglycan (Table 1.3). Pneumococcal autolysis is mediated by these proteins during stationary phase and upon deoxycholate (bile) and β-lactam exposure (López, Ronda & Garcia 1990). Autolysis is also fundamental to cell division (García *et al.* 1999b). The pneumococcus has three autolysins, LytA (*N*-acetyl-muramyl-L-alanine amidase), LytB (glucosaminidase) and LytC (lysozyme), with LytA being the best characterized. All three autolysins have been shown to be important for biofilm formation, as mutants lacking these genes showed up to 50% reduction in biofilm formation and reduced colonization in a rat model (Moscoso, García & López 2006). LytA and LytC are two of the three main lytic enzymes shown to be important in competence-mediated fratricide, along with CbpD (Guiral *et al.* 2005; Eldholm *et al.* 2009). While insertion-inactivation of *lytA* alone had little effect in release of cytoplasmic enzymes following competence activation, double and triple mutants with mutations in *lytC* and *cbpD* had minimal if any competence-mediated cell lysis.

Table 1.3. Biological role, size, and amino acid length of pneumococcal autolysin proteins

	LytA	**LytB**	**LytC**
Biological Role	Acetylmuramidase	Glucosaminidase	Lysozyme
Size (kDa)	36.5	76.4	58.7
Length (aa[a])	318	658	501

[a] aa: amino acid

1.1.5.2.1.1 LytA

LytA cleaves the *N*-acetylmuramyl-L-alanine bond linking the cell wall peptidoglycan and the peptide chain. Autolysin is first synthesized as the inactive form and is activated during stationary phase or by penicillin, and its enzymatic activity is dependent on the presence of choline (Tomasz & Westphal 1971). Mutants with disrupted *lytA* are resistant to lysis in stationary phase, or when exposed to deoxycholate or penicillin (Tomasz, Moreillon & Pozzi 1988). Autolysin contributes to pneumococcal virulence by facilitating the release of pneumolysin and cell wall components that induce pro-inflammatory host responses (Tuomanen *et al.* 1985;

Berry *et al.* 1989; Martner *et al.* 2008). Autolysin-negative mutants have been associated with reduced virulence in numerous animal models. Extended host survival time, rapid bacterial clearance, and increased LD_{50} were observed in infection models using mutated strains (Berry *et al.* 1989; Berry, Paton & Hansman 1992; Canvin *et al.* 1995; Berry & Paton 2000).

1.1.5.2.1.2 LytB and LytC

For LytB and LytC, the choline-binding and the catalytic domains are in the N- and C-termini, respectively, which is opposite to other CBPs (García *et al.* 1999a). LytB is an endo-β-*N*-acetylglucosaminidase responsible for cell wall hydrolysis during cell separation (De Las Rivas *et al.* 2002). Pneumococci with non-functional LytB or LytC result in reduced ability to colonize the nasopharynx of infant rats (Gosink *et al.* 2000).

1.1.5.2.2 *Phosphorylcholine Esterase (Pce)*

Pce (also known as CbpE) is a 69.4 kDa protein with 602 amino acid residues that removes PCho from TAs and LTAs (Vollmer & Tomasz 2001). Insertion duplication mutagenesis of *pce* decreased pneumococcal adherence to Detroit nasopharyngeal cell line *in vitro* and reduction in colonization in an infant rat model (Gosink *et al.* 2000). The activity of Pce in regulating PCho abundance in TA and LTA may modulate ability for pneumococci to adhere to host cells, as well as the ability for C-reactive protein (which binds to pneumococcal PCho) to recognize presence of the organism.

1.1.5.2.3 *Pneumococcal Surface Protein A (PspA)*

The 84-kDa protein impedes complement activation by affecting C3b deposition onto the cell surface (Tu *et al.* 1999). In carriage, PspA interacts with lactoferrin, a host protein thought to exert its bactericidal property by binding to iron, thereby limiting its concentration for bacterial acquisition (Hammerschmidt *et al.* 1999). Shaper *et al.* (Shaper *et al.* 2004) has shown that PspA on the surface as well as extracellular PspA can prevent apolactoferrin (lactoferrin without iron)-mediated bacterial killing, and hypothesized that PspA may block the active site of lactoferrin to prevent the host uptake of iron by the host glycoprotein.

PspA contains an N-terminal signal sequence followed by a region coding for a mature protein with four domains: a positively-charged α-helical N-terminal domain to stabilize the capsular negative charge, a proline-rich domain, a choline-binding domain (CBD), and a hydrophobic C-terminal tail (Yother & Briles 1992). Sequence diversity greater than 20% in *pspA* is found between strains, especially in the α-helical region, possibly as a result of horizontal gene transfer (HGT) (Hollingshead, Becker & Briles 2000). These variations are grouped into three allelic families and six sub-clades (Family 1: clades 1 and 2, family 2: clades 3, 4, and 5, family 3: clade 6). However, homologous and recombinant forms of PspA from multiple strains are able to provide cross-protection in mice against pneumococcal infections by different PspA families and capsular serotypes (Briles *et al.* 2000b; Moreno *et al.* 2010). Sequence variation in the lactoferrin-binding α-helical region between pneumococci is associated with differences in lactoferrin binding (Håkansson *et al.* 2001; Ren *et al.* 2003).

1.1.5.2.4 *Pneumococcal Surface Protein C (PspC)*

PspC is a multifunctional protein known as CbpA (choline-binding protein A), and SpsA (poly Ig receptor-binding protein). PspC contains four domains: the signal leader sequence, N-terminal variable region, a proline-rich sequence, and a C-terminal choline-binding region (CBR). PspC shares homology to PspA in only the proline-rich and C-terminal regions (Hammerschmidt 2007). Similar to *pspA*, *pspC* exhibits extensive sequence diversity between strains, and while most strains contain CBD at its C-terminal, some encode an LPXTG cell wall-binding motif mediating cell anchorage (Iannelli, Oggioni & Pozzi 2002).

PspC is also termed SpsA as it binds to the human secretory IgA via its N-terminal region for transcytosis and invasion across the epithelium (Hammerschmidt *et al.* 1997; Zhang *et al.* 2000). Another name for PspC is CbpA (choline-binding protein A), binding to the laminin receptor in the activated vascular endothelium of the blood-brain barrier to invade into the meninges (Orihuela *et al.* 2009). Additionally, PspC also binds to factor H preventing evasion from the alternative complement pathway (Dave *et al.* 2001). Antiserum of rabbits immunized with PspC cross-react with PspC as well as PspA *in vitro* and in a mouse challenge model (Brooks-Walter, Briles & Hollingshead 1999). The authors hypothesized that cross-reactivity was mediated by

the presence of the proline-rich regions in both proteins (Brooks-Walter, Briles & Hollingshead 1999).

1.1.5.2.5 *Choline-Binding Protein D (CbpD)*

CbpD is a 50-kDa murein hydrolase containing an N-terminal CHAP domain, two SH3 domains, and a C-terminal CBD. Mutants with defective CbpD showed reduced colonization and lung infection in rats and mice, respectively (Gosink *et al.* 2000; Hava 2002). The gene *cbpD* was identified as a late-competence-induced gene (Peterson *et al.* 2004), and CbpD subsequently was found to be involved in competence-mediated fratricide (Chapter 5). CbpD is thought to activate the lytic effects of autolysins LytA and LytC to exert their effects on fratricide (Eldholm *et al.* 2009). The SH3 domains bind to the peptidoglycan, while intriguingly its CBD selectively binds to the poles of target cells; these two interactions were demonstrated to be required for fratricide (Eldholm *et al.* 2010). CbpD, possibly through its CHAP domain may also mediate clumping observed between competent and non-competent cells in the presence of acid, in a yet unknown mechanism related to the release of DNA during fratricide (Håvarstein *et al.* 2006b).

1.1.5.3 Lipoproteins

1.1.5.3.1 *Pneumococcal Surface Adhesin A (PsaA)*

PsaA is a 37-kDa species-specific protein (Russell *et al.* 1990). It was initially postulated to be an adhesin, mainly due to its homology to lipoproteins SsaB and FimA in *Streptococcus sanguinis* and *Streptococcus parasanguinis*, respectively (Sampson *et al.* 1994). Exposure to anti-PsaA antibodies reduced adhesion of pneumococci of different serotypes to host epithelial cells *in vitro* (Romero-Steiner *et al.* 2003). Mutations in *psaA* also showed reduced ability to colonize the nasopharynx and infect the lungs and middle ear of mice, and a 10,000-fold increase in LD_{50} when infected intraperitoneally into mice (Berry & Paton 1996; Marra *et al.* 2002; McAllister *et al.* 2004).

PsaA was subsequently found to be an ABC transporter lipoprotein of the manganese permease complex, encoded by *psaA* found in a locus which also contains genes encoding for an ATP-binding cassette (*psaB*), transmembrane protein (*psaC*), and a thiol peroxidase (*psaD*) (McAllister *et al.* 2004). Mutagenesis of *psa* genes results in

lowered Mn^{2+} uptake and reduced growth *in vitro* unless supplemented with the ion (Marra *et al.* 2002). In addition, PsaA is also suggested to play a crucial role in pneumococcal survival in response to oxidative stress, as mutants lacking PsaA exhibited increased cell death in presence of H_2O_2 compared to wild-type (Tseng *et al.* 2002).

1.1.5.3.2 *Pneumococcal Iron Acquisition and Uptake Transporters (PiaA, PiuA, PitA)*
Three proteins, PiuA (34 kDa), PiaA (42 kDa), and PitA (37 kDa), have been suggested to play roles in iron acquisition and uptake (Brown *et al.* 2001; Brown *et al.* 2002; Whalan *et al.* 2005). These proteins are encoded by three unlinked genes, *piuA* (previously known as *pit1A*), *piaA* (previously known as *pit2A*), and *pitA*. They encode homologues of ATP-binding cassette (ABC) iron transporter of gram-negative organisms and are found in the pneumococci of different serotypes as well as members of *S. mitis* and *S. oralis* (Brown *et al.* 2001; Brown, Gilliland & Holden 2001; Brown *et al.* 2002; Whalan *et al.* 2006). Each of these transporter genes are part of their respective operon each encoding a putative ATPase, the iron receptor, and two genes coding for transmembrane permease proteins. Mutation to one of the three loci results in modest reduction in virulence, while disruption to two or all loci contributed to reduced growth in iron-depleting environment, reduced sensitivity to streptonigrin[13], and attenuation of virulence in a murine pulmonary and systemic model. The *piaA* locus is located in a pathogenicity island resembling a gram-negative source with a reduced G+C content (32.0% compared to ~40.0% in pneumococcal genome) over a length of more than 27 kb, suggesting horizontal acquisition from other species. Immunization with antibodies specific to PiuA and PiaA enhanced opsonophagocytosis by neutrophils *in vitro* and enhanced survival of mice from systemic and lung infections (Brown *et al.* 2001).

1.1.5.4 Peptidoglycan-Associated Proteins

1.1.5.4.1 *Neuraminidase (NanA, NanB, NanC)*
The neuraminidase group comprises of NanA (~115kDa), NanB (~75kDa), and NanC (~82kDa), each containing an N-terminal signal sequence, followed by a lectin-binding domain, and an active site domain. In addition NanA also contains a C-

[13] The activity of streptonigrin is dependent on the presence of iron (Yeowell & White 1982).

terminal LXPTG motif for cell wall adhesion. They cleave sialic acid residues of mucin, glycolipids, glycoproteins, and oligosaccharides. Furthermore, NanA was shown to desialate residues on secretory component, lactoferrin, and IgA2, thereby preventing host clearance and increasing survival on the mucosal surface (King *et al.* 2004). NanA also removes sugar residues of *H. influenzae* and *N. meningitidis*. This is considered as a means of interspecies competition, as removal of these residues confers vulnerability to complement-mediated killing (Shakhnovich, King & Weiser 2002). Most pneumococci contain both *nanA* and *nanB* genes, and *nanC* is present in some isolates (Pettigrew *et al.* 2006). Pneumococcal neuraminidase exhibits intra-species variation, possibly through HGT with viridans streptococci as well as point mutations (King, Whatmore & Dowson 2005). In addition, NanA homologues have been identified in members of *S. mitis* and *S. pseudopneumoniae* (Johnson *et al.* 2010).

Transparent phase variants as well as biofilm cells are associated with increased expression of *nanA* compared to the opaque type (King *et al.* 2004), and biofilm formation was reduced in pneumococci with *nanA* mutation or pneumococci treated with neuraminidase inhibitors. (Parker *et al.* 2009). mRNA levels of *nanA* and *nanB* were also higher in biofilm cells compared to planktonic cells (Oggioni *et al.* 2006).

Pneumococci with defective NanA and NanB resulted in early clearance of pneumococci from the nasopharynx, trachea and lungs, as well as failure to grow in blood in a mouse model of infection following intranasal and intravenous infections (Manco *et al.* 2006). The virulence and pathogenicity of NanC has yet to be elucidated, but Pettigrew *et al.* (Pettigrew *et al.* 2006) reported that pneumococci with *nanC* are more likely to be detected in CSF compared to nasopharynx.

1.1.5.4.2 *IgA1 Protease*

IgA1 protease, encoded by *iga*, is a 200-kDa metalloprotease that cleaves the bond between Pro-227 and Thr-228 in the hinge region of IgA1 (Kilian & Holmgren 1981; Poulsen, Reinholdt & Kilian 1996; Wani *et al.* 1996). The cleavage of IgA1 by the protease enhances adherence to respiratory epithelial cells *in vitro*, probably by altering surface electrochemical properties of the pneumococci (the cationic, hydrophobic properties of F'ab cleaved fragments counteract the anionic, hydrophilic

properties of the pneumococcal capsule) and by enhancing PCho binding to PAFr (Weiser *et al.* 2003). As IgA is the most common immunoglobulin class on mucosal surfaces, and over 90% of IgA is of the IgA1 subclass, the IgA1 protease represents a major colonization factor in the nasopharynx and virulence factor during mucosal infections, allowing the organism to exploit this predominant host immune component for survival.

1.1.5.4.3 *Hyaluronate Lyase (Hyl)*

Hyl is a 107-kDa protein found both anchored to the peptidoglycan and released into extracellular space by *S. pneumoniae*. It cleaves hyaluronic acid, a key glycosaminoglycan component of the extracellular matrix (ECM) of host mucosal surfaces (Berry *et al.* 1994; Hammerschmidt 2007). Hyl contains an N-terminal domain, a carbohydrate adhesion domain, a spacer domain, and the C-terminal LXPTG anchor domain. Fragmented hyaluronan stimulates the production of pro-inflammatory cytokines, thus Hyl activity may play a role in pathogenicity of pneumococcal pneumonia due to the activation and maintenance of the inflammatory response (McKee *et al.* 1996). In addition, Hyl activity may be associated with cases of meningitis, as one study demonstrated the expression of Hyl in all pneumococci isolated from cases of meningitis, compared to 11.5% pneumococci expressing Hyl in the nasopharynx (Kostyukova *et al.* 1995). A recent study has showed that Hyl may allow the pneumococcus to use hyaluronic acid from host as well as respiratory tract colonizers as a carbon source for growth (Marion *et al.* 2012).

1.1.5.4.4 *Pilus Structures*

Pilus-like structures have recently been detected in pneumococci (Barocchi *et al.* 2006; Bagnoli *et al.* 2008; Donati *et al.* 2010) and other members of the Mitis group (Zähner *et al.* 2011). The structures are composed of polymers of LPXTG motifs covalently linked together by the sortase enzyme, which mediate the covalent linkage of LPXTG motif-containing proteins to the cell wall. The genes required for pneumococcal pilus synthesis were mapped to the *rlrA* islet, flanked by *IS*1167 containing seven genes including *rlrA* regulator, the LPXTG-containing proteins

encoded by *rrgA*, *rrgB*, *rrgC*, and the sortase[14] genes *srtB*, *srtC*, and *srtD* (Barocchi *et al.* 2006).

Expression of *rlrA* islet was associated with increased adherence to epithelial cells *in vitro* (Barocchi *et al.* 2006; Nelson *et al.* 2007a), and *rra* mutants were associated with reduction in biofilm formation (Muñoz-Elías, Marcano & Camilli 2008). Mice infected intranasally with mutants lacking RrgA had lower concentration of bacteria in the nasopharynx and higher survival rate compared to wild type strains. A second pilus islet (PI-2) was identified in the pneumococcus (Bagnoli *et al.* 2008), with a similar genetic structure to the *rlrA* islet (PI-1) (Fig. 1.7).

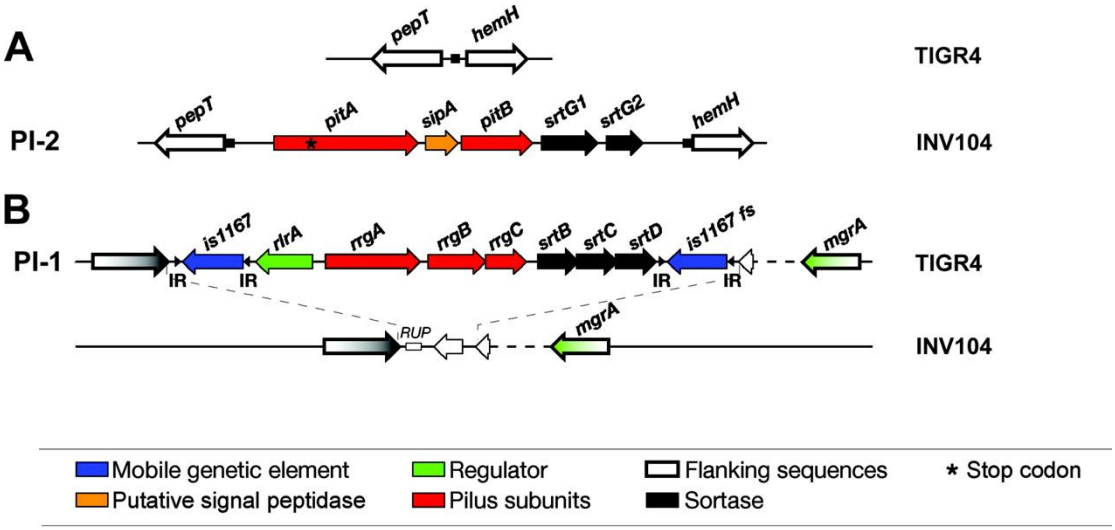


Fig. 1.7. Genetic organization of genes encoding pilus structures (A) PI-2 and (B) PI-1. The functions of proteins/genetic elements encoded by these genes are differentiated by colour groups. The comparison of strains TIGR4 and INV104 highlights the differentiatial distribution of these pili in pneumococci. Comparison of presence/absence of the two pili in additional pneumococci is described in another study (Donati *et al*, 2010). Figure from Bagnoli *et al.* (Bagnoli *et al.* 2008). Journal of Bacteriology authorizes the reuse of journal material including figure shown for academic purposes.

[14] Sortases are transpeptidases which are crucial in the assembly of pili components and has been shown to anchor pili onto cell wall of *Corynebacterium diphtheriae* (Swaminathan *et al.* 2007).

1.1.5.5 Other Surface-Attached Proteins - Pneumococcal Adhesion and Virulence A (PavA)

PavA (62 kDa) is a 551-aa polypeptide mediating pneumococcal binding to host fibronectin, and may play roles in pneumococcal adhesion to host tissues and indirectly mediates virulence. The requirement of PavA in adhesion was clearly demonstrated in cell culture infection studies, where *pavA*-knockout mutants exhibited reduced adherence and invasion of epithelial and endothelial cells (Pracht *et al.* 2005). Also, mice infected with *pavA*-deficient mutants had reduced mortality in models of pneumonia and sepsis (Lau *et al.* 2001; Pracht *et al.* 2005), septicaemia (Holmes *et al.* 2001), and meningitis (Pracht *et al.* 2005). In addition, mutants with deficient *pavA* were cleared from the nasopharynx more rapidly, and were unable to translocate to the bloodstream from the lungs following intranasal challenge (Kadioglu *et al.* 2010).

1.1.5.6 Concluding Remarks for Pneumococcal Proteins

While the capsule has long been recognized as a major virulence factor of the pneumococcus, additional proteins, whether cytosolic, extracellular, or surface-associated, also contribute to virulence and pathogenicity. Furthermore, the majority of these factors have important roles in colonization. As the human nasopharynx is the most predominant ecological niche, it is possible that these factors have evolved and adapted to better adapt during colonization, and in relatively rare cases of infection, these factors also mediate virulence in order to survive in the host.

1.2 Pneumococcal Diseases

The World Health Organization estimates that pneumococcal diseases account for 1.6 million deaths worldwide per year, with the majority of which occur in children under the age of 5 years and the elderly (World Health Organization 2007). The ten countries with the highest mortality rates from invasive pneumococcal diseases (all in Asia and Africa) constitute over 60% of all pneumococcal deaths in children (O'Brien *et al.* 2009). The pneumococcus commonly causes relatively mild diseases such as otitis media and sinusitis, and more severe clinical manifestations such as pneumonia. The pneumococcus also causes infections in normally-sterile sites, leading to invasive diseases such as septicaemia and meningitis. In addition, the organism occasionally causes pyogenic arthritis, osteomyelitis, pyomyositis, necrotizing faciitis, endocarditis, pericarditis, peritonitis, paritotis, endophthalmitis, and conjunctivitis (Butler 2004). Pneumococcal diseases are usually caused by a small number of serotypes; around 20 serotypes cause 70-90% of IPD worldwide, and serotypes 1, 5, 6A, 6B, 14, 19F, and 23F collectively are responsible for at least 50% of IPD in children < 5 years in every continent (Örtqvist 2001; Johnson *et al.* 2010). A higher number of serotypes are responsible for an equal proportion of IPD in Africa and Asia than in North America, suggesting a greater diversity of serotypes causing IPD in the former two continents (Johnson *et al.* 2010).

S. pneumoniae enters the bloodstream from the lungs causing septicaemia. Activation of host complement system promotes transcytosis through the blood-brain barrier and subsequent pneumococcal replication in the brain. Meningitis can also occur in the absence of apparent bacteraemia and pneumococcus-positive blood culture; it is hypothesized that other infection routes allow bypassing of the bloodstream to cause meningitis without septicaemia (Marra & Brigham 2001; Filippidis & Fountas 2009).

1.2.1 Conjunctivitis

Outbreaks of conjunctivitis frequently involve acapsulate and non-serotypable (NT) pneumococci (Shayegani *et al.* 1982; Smart, Dougall & Girdwood 1987; Carvalho *et al.* 2003; Crum *et al.* 2004; Haas *et al.* 2011). Genetic analyses show that most pneumococcal conjunctivitis is caused by clonal complexes of related NT strains, including clonal complex of ST344, and are genetically diverse based on PFGE and MLST (Martin *et al.* 2003; Crum *et al.* 2004; Buck *et al.* 2006; Porat *et al.* 2006).

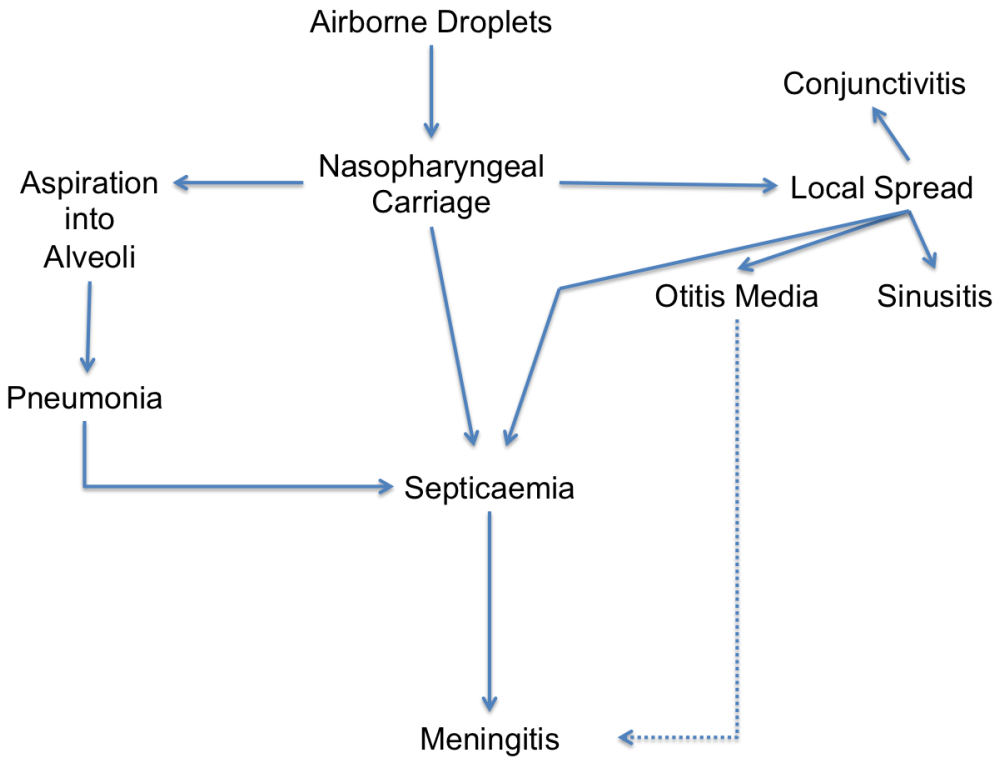


Fig. 1.8. Pneumococcal diseases and pathogenic route. Pneumococci enter host via airborne droplets into the nasopharynx, where they colonize. Local spread from the nasopharynx into the middle ear, sinus, and conjunctiva leads to otitis media, sinusitis, and conjunctivitis, respectively. From the nasopharynx, pneumococci can also be aspired into alveoli leading to pneumonia. Septicaemia can be caused by translocation of pneumococci into blood. Further translocation from the blood to the meninges can cause meningitis, however some studies hypothesize that pneumococci can cause meningitis without septicaemia via other infection routes. Figure modified from Bogaert, de Groot and Hermans (Bogaert, de Groot & Hermans 2004).

Some of the NT STs from the conjunctiva are also distinct from NT isolated in other anatomical sites. This suggests that loss of capsule alone does not explain these NT strains' likelihood to cause conjunctivitis, and that additional genetic factors associated with these distinct STs may play contributing roles in these NT strains (Berrón *et al.* 2005).

1.2.2 Otitis Media (OM)

OM is the leading cause for visits to doctors among children. Pneumococcus is among the leading bacterium responsible for otitis media in both adults and children, followed by *H. influenzae* and *M. catarrhalis*. High-density population settings such

as crowded homes and day-care centres have been shown to be a risk factor development of acute OM in child populations (Faden *et al.* 1997; Paradise *et al.* 1997). Otitis media may be more common in indigenous populations such as Australian Aboriginals and Native Americans (Leach 1999; Curns *et al.* 2002). OM is the most common infection in children and diagnosis for antibiotic prescription. This leads to the rapid emergence of antibiotic-resistant pneumococci worldwide, including strains of serotypes not directly targeted by the heptavalent pneumococcal conjugate vaccine (Section 1.4.2) (Huebner *et al.* 2003; Xu *et al.* 2009; Porat *et al.* 2010). Middle ear infections in humans (Hall-Stoodley *et al.* 2006; Hoa *et al.* 2009b) and chinchillas (Ehrlich *et al.* 2002; Hoa *et al.* 2009a; Reid *et al.* 2009; Weimer *et al.* 2010) have been associated with the formation of pneumococcal biofilms (Section 1.5.9).

1.2.3 Pneumonia

Pneumonia accounts for nearly 20% of childhood deaths worldwide, and 50% of pneumonia cases can be attributed to the pneumococcus, with a case-fatality rate of over 10% in Africa (O'Brien *et al.* 2009; Jambo *et al.* 2010). In addition to children, the elderly and the immunocompromised are also at heightened risk. Pneumococcal pneumonia is triggered by the interaction between pneumococcus and the defense mechanisms of the lungs, following inhalation of pneumococci into alveolar air spaces. Alveolar macrophages are normally efficient in clearing the infection by opsonophagocytosis (Gordon *et al.* 2000). Failure to clear the infection results in rapid bacterial proliferation in the alveoli. At high bacterial concentration, alveolar macrophages are overwhelmed and phagocytic activity is lowered, with concomitant activation of pro-inflammatory cytokines. The development of pneumococcal pneumonia in CD1 mice is described in detail in Bergeron *et al*. (Bergeron *et al.* 1998).

1.2.4 Septicaemia/Bacteraemia

Despite the availability of antibiotics and vaccines, the case-fatality rate of pneumococcal bacteremia remains at approximately 20% (Balakrishnan *et al.* 2000). Suspected Cases of bacteremia are usually treated with amoxicillin and third-generation cephalosporin, however the recent isolation of a serotype 19A strain highly resistant to penicillin and ceftriaxone is alarming (Shouval *et al.* 2010). Nearly 50%

of bacteremia fatalities occur at the early onset, therefore rapid diagnosis of bacteremia by molecular detection directly on blood samples provides an efficient alternative to the conventional blood culture method which is time-consuming (Paolucci, Landini & Sambri 2010). Recently, Kadioglu *et al.* (Kadioglu *et al.* 2011) demonstrated that male mice were more susceptible to sepsis with shorter survival time and a more pronounced inflammatory response.

1.2.5 Meningitis

The pneumococcus is a leading bacterial agent of meningitis in children, with a high disease incidence during the dry season (January and February) in the African pneumococcal meningitis belt in the sub-Sahara region, with a high proportion of cases caused by serotype 1 (Leimkugel *et al.* 2005; Yaro *et al.* 2006; Gessner, Mueller & Yaro 2010). Mortality rates for meningitis may be as high as nearly 80%, and survivors suffer from neurological sequelae such as hearing loss, learning deficits, mental retardation, sensory motor deficits, and seizure disorders (Grimwood *et al.* 1995; Marra and Brigham 2001). These outcomes are probably due to tissue damage inflicted in the brain during the inflammatory process.

Meningitis occurs by the breaching the blood-brain barrier (BBB) and transcytosis across into the subarachnoid space. Initial attachment to the brain capillary endothelial cells (BMEC) requires CbpA, and the upregulation of host PAFr is required for entrance into the subarachnoid space (Ring, Weiser & Tuomanen 1998). Increase in PAFr expression is mediated by TNF-α production stimulated by the pneumococcal cell wall (Cundell *et al.* 1995; Freyer *et al.* 1999). Once inside the CNS, bacterial replication occurs and inflammation is induced in this sterile environment to cause the pathogenic traits associated with meningitis.

1.3 Antibiotic Treatment and Resistance

1.3.1 β-Lactams

Twenty years after the introduction of penicillin in 1940, penicillin-nonsusceptible pneumococci were first isolated from the sputum of an adult in Australia (Hansman & Bullen 1967), and Papua New Guinea (Hansman *et al.* 1974). Detection of resistant isolates from children with systemic diseases in South Africa was also recorded (Appelbaum *et al.* 1977). By the 1990s the majority of pneumococci isolated in studies in Spain, Hungary, and South Africa were resistant to penicillin and other β-lactams, some of which with MICs 1000-fold of susceptible strains (Fenoll *et al.* 1991; Marton *et al.* 1991; Liñares *et al.* 1992).

The rise in penicillin resistance was associated with the increase in use of β-lactams in the early 1990s (Granizo *et al.* 2000). The emergence of penicillin-nonsusceptible pneumococci can be attributed to expansion of resistant clones, many of which characterized by the Pneumococcal Molecular Epidemiology Network (Mcgee *et al.* 2001) (Fig. 1.9). These resistant clones are associated with serotypes commonly found in carriage such as members of serogroups 6 and 19. Conversely, invasive serotypes but ones that are rarely detected in carriage (serotype 1) are rarely resistant, even in countries where penicillin nonsusceptibility is common, consistent with the hypothesis that nonsusceptibility is mainly acquired and transmitted during carriage (Yagupsky *et al.* 1998; Marimon *et al.* 2009; Fenoll *et al.* 2010).

1.3.2 Mechanism of β-Lactam Resistance

Genes encoding β-lactamases have yet to be identified in the pneumococci. Non-susceptibility to β-lactam in pneumococci is due to the reduced affinity of the drug to the six penicillin-binding proteins (PBPs). These proteins are responsible for the cross-linking of peptidoglycan strains. β-lactams bind to the active site region of these proteins, inactivating peptidoglycan linking. β-lactam resistance is mediated by genetic changes of PBP genes resulting in the reduced affinities of PBPs to the drugs. PBPs are classified into 3 groups: high-molecular weight (HMW) class-A proteins PBP1a, PBP1b, and PBP2a; HMW class-B proteins PBP2x and PBP2b; and a low-molecular weight PBP3. Mutations in *pbp2b* and/or *pbp2x* confer low-level β-lactam resistance, while alterations in *pbp1a* contribute to higher level of resistance if mutations in *pbp2b* and/or *pbp2x* are present (Smith & Klugman 1998).

Fig. 1.9. Spread of penicillin-nonsusceptible *S. pneumoniae*. A total of 26 pneumococcal clones non-susceptible to penicillin and other drugs have been characterized by the PMEN. These clones are named according to the country where they were first identified, the serotype (superscript), and the PMEN clone number in chronological order. These clones are responsible to the spread of drug resistance due to their expansion.

Resistant isolates contain major sequence blocks with extensive variation in *pbp* genes. Sequence variation of over 20% between resistant and sensitive strains can be observed for *pbp1a* (Martin, Sibold & Hakenbeck 1992), *pbp2x* (Laible, Spratt & Hakenbeck 1991), and *pbp2b* (Dowson *et al.* 1989). These variable sequences are likely to have been acquired from the related *S. mitis* and *S. oralis* (Dowson *et al.* 1989; Dowson *et al.* 1993; Sibold *et al.* 1994). Recombination between *pbp* genes can occur within the *pbp* genes or in the flanking regions, as in the case of *ddl* hitchhiking effect with *pbp2b* and the transformation of *cps* along with *pbp2x* and *pbp1a* (Coffey *et al.* 1999; Enright & Spratt 1999; Trzciński, Thompson & Lipsitch 2004; Brueggemann *et al.* 2007).

Amino acid alterations at the penicillin-targeting active site contribute to changes in the peptidoglycan structure. Resistant pneumococci contain a higher proportion of hydrophobic branched muropeptides in their cross-linkages, compared to susceptible pneumococci containing mainly linear stem peptides (Garcia-Bustos & Tomasz 1990). A possible explanation for its role in resistance may be that branched peptides are more efficient competitors to the PBPs than β-lactams. This suggests that changes in penicillin-binding proteins (PBPs) as a result of reduced drug affinity also change substrate specificity in murein peptide chain synthesis (Garcia-Bustos & Tomasz 1990).

1.3.3 Macrolides

Macrolides inhibit translation by binding to the 23S rRNA of the 50S bacterial ribosome subunit. Resistance to macrolides involves 23S rRNA methylation (mainly mediated by *erm(*A), *erm*(B)), and efflux pump (*mef* genes). Both classes of genes can be found flanked by genes encoding transposable elements (McDougal *et al.* 1998; Gay & Stephens 2001; Croucher *et al.* 2011). The increase of macrolide resistance in some areas may be correlated by increase in macrolide usage (Karlowsky *et al.* 2009) selecting for survival and expansion of strains containing these genetic elements. Also, strains containing both *mef*(A) and *erm*(B) are associated with multi-drug resistance, and are also hyper-recombinant (Lee, Song & Ko 2010). Depending on geographical region, macrolide nonsusceptibility rates may be as low as 1% in Mozambique to over 90% in Taiwan (Van Bambeke *et al.* 2007).

1.3.4 Fluoroquinolones

The quinolone is a broad class of antibiotic drugs affecting DNA replication. Most quinolones belong to the fluoroquinolone group, containing fluorine in the central carbon ring. Quinolones and fluoroquinolones target the GyrA and ParC proteins of the DNA gyrase and topoisomerase IV proteins, respectively, which are involved in regulating DNA supercoiling. Fluoroquinolones halt the replication fork during DNA replication. DNA gyrase is a heterodimer consisting of GyrA and GyrB, and the topoisomerases IV is a heterodimer of ParC and ParE. Resistance to fluoroquinolones is associated with amino acid substitutions in GyrA, ParC, and ParE (Tankovic *et al.* 1996; Perichon, Tankovic & Courvalin 1997). ParC appears to be the primary target for quinolones in *S. pneumoniae*, while additional mutations in *gyrA* results in higher resistance. Mutations in *parE* and *gyrB* may also be present, but are not found in all strains with quinolone resistance (Bast *et al.* 2000; Richter *et al.* 2005). The majority of spread is as a result of clonal expansion of nonsusceptible strains, with a small proportion from interspecies transfer (Balsalobre *et al.* 2003; Stanhope *et al.* 2005). High nonsusceptibility rates have been documented in Hong Kong, Sri Lanka, the Philippines, and Korea (Song *et al.* 2004).

1.3.5 Trimethoprim-Sulfamethoxazole (Co-trimoxazole, SXT)

SXT is mainly used as a first-line antibiotic in developing countries for dual-treatment for malaria and pneumonia, mainly because of its low cost and its effectiveness as a daily prophylaxis for HIV/AIDS patients (Wiktor *et al.* 1999; Mermin *et al.* 2004; Hamel *et al.* 2008). SXT targets dihydrofolate reductase (DHFR) encoded by *folA* and dihydropteroate synthase (DHPS) encoded by *folP*, respectively. DHFR and DHPS are enzymes involved in the folate biosynthesis pathway, required for purines and amino acid synthesis.

SXT-nonsusceptible pneumococcus was first isolated from a patient with chronic bronchitis in 1972 (Howe & Wilson 1972). Since then, resistance has increased rapidly, with nonsusceptibility prevalence well above 50% across continents (Joloba *et al.* 2001; Arifeen *et al.* 2009). In the laboratory, single amino acid substitutions has been shown to mediate non-susceptibility to trimethoprim, namely the leucine substitution from isoleucine in amino acid position 100. However, clinical isolates of resistance pneumococci are usually associated with multiple substitutions, along with

I100L (Adrian & Klugman 1997; Wilén *et al.* 2009). These additional alterations, possibly mediated through mutations or interspecies recombination, may moderate affinity of enzyme to the natural substrate in response to the reduced susceptibility to the antibiotics.

1.4 Prophylaxis – Pneumococcal Vaccines

The development of pneumococcal vaccines was prompted by 1) the high mortality and morbidity associated with pneumococcal infections despite antibiotic treatment, and 2) the rapid emergence of antibiotic non-susceptible pneumococci. Four pneumococcal vaccines have been licensed and more are in development. Current licensed vaccines target the pneumococcal polysaccharide capsule, providing protection from infection of the most commonly encountered serotypes. The following sections will describe the 23-valent polysaccharide (PPV23), seven-valent (PCV7), and 13-valent (PCV13) conjugate vaccines (Table 1.4). A nine-valent (PCV9) vaccine (containing PCV7 serotypes plus 1 and 5) was also effective in reducing IPD in a clinical trial in the Gambia (Cutts *et al.* 2005), however it was not developed further because of the expanded valency provided by PCV13.

Table 1.4. Currently licensed pneumococcal vaccines

Vaccine	Serotypes Covered	Year Licensed
PPV23 (Pneumovax)	1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19F, 19A, 20, 22F, 23F, 33F	1983
PCV7 (Prevnar)	4, 6B, 9V, 14, 18C, 19F, 23F	2000
PCV10 (Synflorix)	1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, 23F	2009
PCV13 (Prevnar 13)	1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F	2010

1.4.1 PPV23

The 23-valent pneumococcal polysaccharide vaccine (PPV23) was designed to reduce invasive diseases and pneumonia in adults. The vaccine, formulated by Merck, containing 25 µg of purified capsular polysaccharide from each of the included serotypes. These serotypes covered over 90% and 60% of isolates responsible for

invasive diseases in the US in 2000 (Rudolph, 2000). PPV23 is recommended for the healthy individuals over the age of 65, and high-risk individuals over the age of 2, including asplenic patients, people with chronic renal, heart, lung, liver diseases, diabetes mellitus, and the immunocompromised (Melegaro & Edmunds 2004).

The pneumococcal capsule is a T-cell-independent type 2 (TI-2) antigen, binding directly to B-cells via its repeating sugar epitopes. Immune response to TI-2 requires co-stimulation with CD21, which is expressed at low levels in infants (Rijkers *et al.* 1998). PPV23 are thus not immunogenic in children under the age of 2. Antibody responses following PPV23 immunization may be transient in other age groups. Repeated administration of PPV23 may also cause immune hyporesponsiveness, and this state of immune tolerance may stem from memory B cells being exposed to excessive amounts of polysaccharide antigens (O'Brien, Hochman & Goldblatt 2007). PPV23 also provides limited protection in mucosal infections such as AOM in children and pneumonia in the elderly (Douglas & Miles 1984; Jackson *et al.* 2003). These reports demonstrate the sub-optimal efficacy of PPV23 for age groups most vulnerable to pneumococcal infections, possibly due to its polysaccharide nature. Therefore, the heptavalent pneumococcal conjugated vaccine (PCV7) was introduced in an attempt to shift the capsule-targeted immune response to an earlier age (see below).

1.4.2 PCV7

Polysaccharide antigens from each of the seven serotypes in PCV7 are individually conjugated to a non-lethal diphtheria toxoid protein (cross-reactive material 197, CRM197), which converts the polysaccharide antigen to a T-dependent response present in young children. The PCV7 was first introduced in 2000 in the United States, recommended for all children aged 2-20 months as well as those aged 2-5 years with increased risks. By 2008, PCV7 had been licensed for use in more than ninety countries. In addition, numerous developing countries have recently been approved for the use of PCV7 (World Health Organization 2008).

In the US, PCV7 administration was associated with over 97% reduction in IPDs caused by vaccine serotypes (VTs) and 89.1% reduction in IPDs overall in infants who received PCV7 (Black *et al.* 2000). Similar but more modest reductions observed

in individuals of other age groups as a result of herd immunity[15] (Whitney *et al.* 2003; Carter 2006; Miller *et al.* 2011). Use of PCV7 has also contributed to the decrease of IPD caused by antibiotic-nonsusceptible pneumococci in children and adults (Black *et al.* 2004; Hampton *et al.* 2011). Decreases in IPD by VTs were also seen in high-risk groups, such as those over the age of 65, indigenous populations, and HIV-positive infants (Klugman *et al.* 2003; O'Brien *et al.* 2003; Lexau *et al.* 2005). Studies also documented more modest reductions in otitis media (Eskola *et al.* 2001; Fireman *et al.* 2003), and pneumonia (Black *et al.* 2002) cases.

PCV7 was associated with reductions in colonization by VT, with VT carriage rates as low as 3% after PCV implementation (Dagan *et al.* 2002; Huang *et al.* 2009). The protective property of PCV in carriage extends to vaccine-related serotypes; serotype 6A carriage rate was significantly lower in toddlers who received the nonavalent pneumococcal conjugate vaccine (PCV9, contains serotypes in PCV7 in addition to serotypes 1 and 5) targeting serotype 6B. The carriage of antibiotic-resistant pneumococci was also lowered upon the administration of conjugate vaccines in children, as the majority of nonsusceptible strains are of serotypes included in PCV7 (Dagan *et al.* 2003).

1.4.3 Serotype Replacement

Due to limited serotype protective coverage, conjugated vaccines enable serotypes not covered in the vaccine to fill up niches previously occupied by VTs, a phenomenon termed serotype replacement. Since PCV7 use, serotype 19A[16] is among the most prevalent NVT to cause IPD in the post-vaccine era across continents, with serotypes 15B/C, 35B, 22F, 6C, 7F, and 11A also common (Singleton *et al.* 2007; Carvalho *et al.* 2009; Hsu *et al.* 2009; Huang *et al.* 2009; Hanage *et al.* 2010; Hanquet *et al.* 2010; Gladstone *et al.* 2011a; Miller *et al.* 2011; Wroe *et al.* 2011). Replacements have also been observed in carriage (Pelton, Loughlin & Marchant 2004; Nahm *et al.* 2009; Tocheva *et al.* 2010; Tocheva *et al.* 2011) and AOM (McEllistrem *et al.* 2003). Rates of antibiotic resistance in replacement serotypes are also increasing (Hanage *et al.* 2007; Hicks *et al.* 2007; Simões *et al.* 2011a). Serotype replacement could be

[15] Herd immunity: protection of non-vaccinated individuals due to the large proportion of vaccinated individuals in that transmission of disease-causing organisms is interrupted.

[16] PCV7 targeting serotype 19F does not provide cross-protection for serotype 19A (Dagan *et al.* 2002)

mediated by the expansion of pre-existing NVT clones and clonal complexes, such as that of ST199 and ST320, both in serotype 19A (Hanage *et al.* 2007; Gladstone *et al.* 2011a), and serotype switching, where the *cps* locus encoding a NVT capsule is acquired by a virulent strain of a VT, generating vaccine escape recombinants (see below).

The clinical impact of serotype replacement depends on the invasiveness of the serotype that is replacing. Host factors, transmission environment (crowded vs. non-crowded), and vaccination coverage within the host population also contribute to serotype replacement as these are variables likely to affect pneumococcal transmission between hosts (Long 2005). It must be noted that temporal fluctuations in serotype prevalence may take place in the absence of PCVs, as seen carriage in Icelandic children (Tomasson, Gudnason & Kristinsson 2005) and IPD in Scotland (Jefferies *et al.* 2010a). However, serotype replacement could be an artifact of unmasking[17] within multiple colonization (Azzari *et al.* 2008). The effects of serotype replacement can be assessed more accurately by using sensitive serotyping methods (Chapters 3 and 4).

1.4.4 Serotype Switching

Serotype switching is the horizontal acquisition of the *cps* operon, leading to the expression of a different capsule. The change of serotype may involve parts of the operon (Park *et al.* 2007) or the entire operon, sometimes including adjacent *pbp* genes (Coffey *et al.* 1998; Nesin, Ramirez & Tomasz 1998; Brueggemann *et al.* 2007; Pillai, Shahinas & Buzina 2009; Croucher *et al.* 2011; Golubchik *et al.* 2012). Serotype switching has also been seen in laboratory settings (Kelly, Dillard & Yother 1994; Trzciński, Thompson & Lipsitch 2004). Other studies infer capsular switching as comparison of two strains of different serotypes having identical or highly similar genetic properties determined by PFGE, MLST, and antibiotic-resistance and virulence genes (Barnes, Whittier & Gilligan 1995; Jefferies *et al.* 2004; Moore *et al.* 2008; Hanage *et al.* 2011). As different capsule types have varying disease potentials (Kelly, Dillard & Yother 1994; Brueggemann *et al.* 2003; Sandgren *et al.* 2004), serotype switching may significantly change the invasiveness of a pneumococcus. It

[17] Unmasking: NVT may not have increased in prevalence, but are more frequently detected simply due to the reduction of VT in carriage.

reduces the efficacies of pneumococcal vaccines as recombinants are able to escape serotype-specific immune responses. Comparison of serotype, genetic background (by MLST), and *pbp* genes on over 200 international clones provides evidence that serotype switching is relatively common compared to horizontal transfer of penicillin-resistance determinants (Stanhope *et al.* 2007).

1.4.5 PCV10

GlaxoSmithKline introduced a 10-valent conjugate vaccine (PCV10) in 2009, containing serotypes included in PCV7 in addition to 1, 5, and 7F. Two of the serotypes are conjugated to either a tetanus (18C) or a diphtheria (19F) toxoid, while the remaining eight serotypes are conjugated to the cell surface protein D of *H. influenzae*. This is thought to confer protection against AOM caused by nontypable *H. influenzae* in addition to IPD (Gladstone *et al.* 2011).

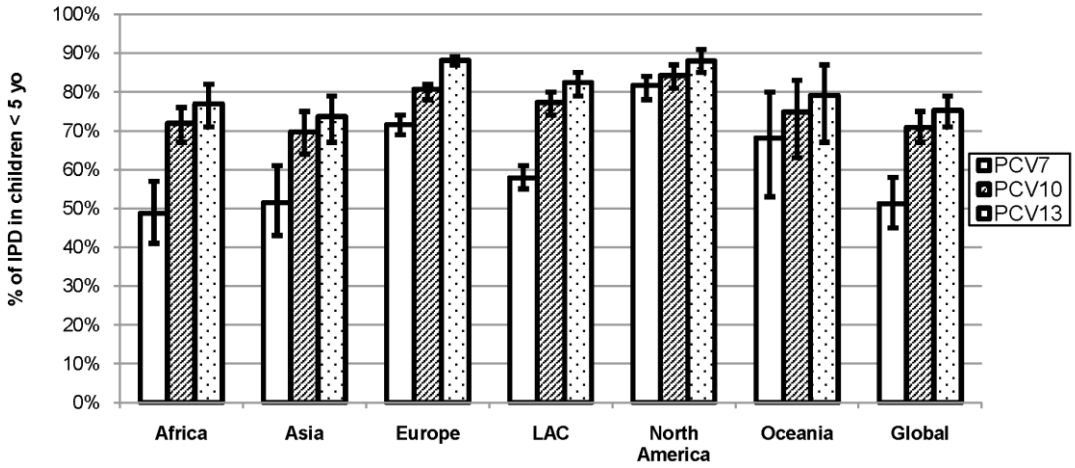
1.4.6 PCV13

A 13-valent conjugate vaccine was licensed in 2009 in Chile and 2010 in the United States and Europe, covering PCV7 serotypes in addition to serotypes 1, 3, 5, 6A, 7F, and 19A (Reviewed in Jefferies *et al.* 2011). It is estimated that the global vaccine coverage of PCV13 exceeds 70% of IPD in children < 5 years, ranging from 74% to 88% depending on geographical region (Johnson *et al.* 2010), and the coverage is likely to be higher given its cross-protective properties to serotypes 6C and 7A (Cooper *et al.* 2011). Subjects immunized with PCV13 showed higher levels of functional IgG (as assessed by opsonophagocytosis assay [18], OPA) than those immunized with PPV23, for at least 12 of the 13 serotypes, except for serotype 4 (Scott *et al.* 2007). For serotypes included in PCV7, infant IgG levels were similar to that elicited by PCV7, in addition to fold-increases in levels of IgG targeting the extra six serotypes not covered in PCV7 (Bryant *et al.* 2010). The Global Alliance for Vaccines and Immunization (GAVI) aimed to provide easier access to PCV13 in resource-poor nations. As of 2011, PCV13 has been introduced in developing countries of three continents (Nicaragua, Guyana, Yemen, Sierra Leone, Kenya, Mali,

[18] The OPA analyze efficacy of vaccine by assessing antibody's ability to opsonize bacteria. This is an accompaniment to ELISA detection of IgG concentration, because IgG concentration alone does not equate presence of functional antibody and thus protection, as seen in the elderly (Romero-Steiner *et al.* 1997).

Democratic Republic of Congo, Benin, Cameroon, Rwanda, Burundi, Ethiopia, Malawi) (Global Alliance for Vaccine and Immunization 2012).

Fig. 1.10. PCV7, PCV10, and PCV13 coverages of serotypes causing IPD in children under the age of 5 by continent. Figure taken from Johnson *et al.* (Johnson *et al.* 2010). PLoS grants unrestricted use of contents including figure above for academic purposes.



1.4.7 Protein-Based Vaccines

The lack of efficacy of PPV23 in infants and the limited serotype coverage of conjugate vaccines call for the design of protein-based vaccines. Such vaccines should be immunogenic to children under the age of 2, provide protection from pneumococci of different serotypes and strains, and reduce cost for its widespread use in developing countries.

Protein antigens can be used either as purified protein vaccine or as an adjuvant of polysaccharide conjugate vaccines. Potential antigen targets include a purified, non-cytotoxic variant of Ply (pneumolysoid) that provided protection in mice against pneumonia and septicaemia across different serotypes (Alexander *et al.* 1994; Briles *et al.* 2003; Kirkham *et al.* 2006; Harvey *et al.* 2011). Despite the antigenic variability, administration of PspA subcutaneously, orally, and nasally, has also been shown to confer protection to sepsis and carriage in mice and human challenged with pneumococci of different serotypes and PspA types (Talkington *et al.* 1991; Wu, Nahm & Guo 1997; Yamamoto *et al.* 1997; McCool *et al.* 2002). The mixture of pneumolysoid and PspA immunization may also provide additive protection (Briles *et al.* 2003). For protection against pneumococcal carriage, it has been shown that PsaA

immunization is superior to PspA and pneumolysoid, with the mixture of PsaA and PspA providing the best protection against carriage (Briles *et al.* 2000a). Other potential vaccine targets include PspC, LytA, PiuA, PiaA, Hyl, and IgA1 protease.

1.5 Pneumococcal Colonization

The human nasopharynx is the predominant ecological niche for the pneumococcus, and rarely is the organism found in animals (Whatmore *et al.* 1999; Chi *et al.* 2007; Denapaite & Hakenbeck 2011). Colonization is required for pneumococcal disease as well as dissemination to other hosts. Local spread of pneumococci from the URT can result in otitis media and sinusitis, while pneumococci entry into alveolar spaces contributes to the development of pneumonia (Reviewed in Bogaert, de Groot & Hermans 2004). Serotyping and genetic typing methods reveal that serotypes and strains causing diseases overlap those detected in carriage, suggesting that such strains colonize the host before causing diseases, but strains colonizing are more genetically diverse compared to invasive pneumococci (Gray, Converse & Dillon 1980; Müller-Graf *et al.* 1999; Robinson *et al.* 2001; Brueggemann *et al.* 2003). Of the pneumococci that cause diseases, colonization of the nasopharynx is followed by via receptor-mediated uptake and transcytosis, and the entry of submucosal layers (Cundell *et al.* 1995).

1.5.1 Mechanism of Colonization

Host surface proteins are involved in pneumococcus-host interaction. CBPs may be involved in hydrophobic and/or electrostatic interaction with epithelial cells (Swiatlo *et al.* 2002). PspA binds to non-inflamed, resting epithelial cells via its Glc*N*Ac receptors, which may be exposed following cleavage of glycolipids, glycoproteins, and oligosaccharides by NanA. Local inflammatory response may alter the expression of host platelet-activating factor receptor (PAFr). PCho on TA and LTA acts as an anchor for a number of CBPs, but PCho itself may also act as an adhesin, binding to host PAFr, facilitating transcytosis during progression from carriage to invasion as well as crossing the blood-brain barrier (Cundell *et al.* 1995; Ring, Weiser & Tuomanen 1998). Furthermore, CbpA has been demonstrated to bind to sialic acid and lacto-N-neotetraose, known pneumococcal ligands for cytokine-activated epithelial host cells (Rosenow *et al.* 1997). NanA can also alter surface structures of

competing co-colonizing bacteria (Section 1.5.7). The result of this is the exposure of adhesion factors, modulation of host clearance glycoproteins, reduction of competitiveness of co-colonizing species, and a supply of carbon source for growth (Tong *et al.* 1999; Shakhnovich, King & Weiser 2002; King *et al.* 2004; King, Hippe & Weiser 2006).

1.5.2 Age-Related Epidemiology in Carriage

All individuals are likely to have been colonized by pneumococci at least once, and most by two years of age (Gray, Converse & Dillon 1980). Pneumococcal colonization is most common in infants under the age of six. Colonization may begin as early as four days after birth (Gray, Converse & Dillon 1980), and prevalence peaks at the age of one to three years (Bogaert, de Groot & Hermans 2004). Children in day care centers have high carriage rates, possibly due to the close contact facilitating transmission (Bogaert *et al.* 2001; Regev-Yochay *et al.* 2004b; Sá-Leão *et al.* 2008). Colonization prevalence as well as duration of carriage are lower in older children and adults (Gray, Converse & Dillon 1980; Hansman *et al.* 1985; Regev-Yochay *et al.* 2004b; Hussain *et al.* 2005; Chen *et al.* 2007; Hogberg *et al.* 2007; Abdullahi *et al.* 2008). In adults, carriage rates are higher in areas of close contacts, such as prisons, barracks, dormitories, and nursing homes. Carriage rates in adults tend to be low, but may be higher in special populations, disease outbreaks, or when living with young children (Grennan *et al.* ; Nuorti *et al.* 1998; Flamaing *et al.* 2010).

1.5.3 Epidemiological Differences in Serotypes during Colonization

Different serotypes vary in colonization prevalence and duration, and while the chemical properties of the capsule may account for such difference, host age, geographical location, cohort type, and other factors also play roles in shaping this prevalence.

1.5.3.1 Age-Related Differences

Disparities in prevalent carriage serotypes between different age groups are best described when analyzing cross-sectional studies of numerous age groups within a single community (Regev-Yochay *et al.* 2004b; Hussain *et al.* 2005; Hammitt *et al.* 2006; Abdullahi *et al.* 2008). While serotypes such as 6B, 6A, and 14 are common in children and adults, prevalence of PCV7 serotypes as a group decreases with

increasing age. PCV7, in addition to serotypes 6A, 35B, and 23A, are commonly found in young children. Prevalence of colonizing PCV7 serotypes decreases in older children, and serotypes 3 and 8 and groups 10, 11, 16, 22, and NT serotypes increase in prevalence by age until adulthood (Bogaert *et al.* 2001; Regev-Yochay *et al.* 2004b).

1.5.3.2 Geographical Distribution of Different Serotypes

In general serotypes of the PCV7 are commonly detected across the globe. Serotypes 6B, 19F, and 23F are among the most prevalent colonizing serotypes across continents (Bogaert *et al.* 2006; Laval *et al.* 2006). In the pre-PCV7 era, high PCV7 coverage rates are seen in Europe and North America, and vaccine coverage rates may be lower in Asia, Middle East, and Africa (Dagan *et al.* 1996; Soewignjo *et al.* 2001; Hill *et al.* 2008), as well as in special populations such as the American Natives and Australian Aboriginals (Mackenzie *et al.* 2007; Millar *et al.* 2009; Mackenzie *et al.* 2010). In addition, prevalence of different serotypes can be seen across different countries within a continent (Bogaert, de Groot & Hermans 2004), or different cohorts within a country or city (Givon-Lavi *et al.* 1999; Hill *et al.* 2006).

1.5.4 Risk Factors for Colonization

1.5.4.1 Crowding

Crowded environments represent a reservoir for the rapid transmission of pneumococci between individuals in close contact, and pneumococcal disease outbreaks tend to occur in these environments. Day care center attendees were associated with high carriage rates (Rosén *et al.* 1984; Dagan *et al.* 1996; Vives *et al.* 1997; Bogaert *et al.* 2001; Peerbooms & Engelen 2002; Vestrheim *et al.* 2008a). Children are potentially the source of pneumococci in a family setting where transmission to mothers and infants occur. The high prescription of antibiotics in young children provides a selective pressure for colonization and dissemination of antibiotic resistant strains in day care centers (Barnes, Whittier & Gilligan 1995). A large-scale longitudinal carriage study conducted in Israeli day-care center children revealed pneumococcus-positivity in 77% of cultures, with over half of the isolates being resistant to at least one antibiotic (Givon-Lavi *et al.* 1999).

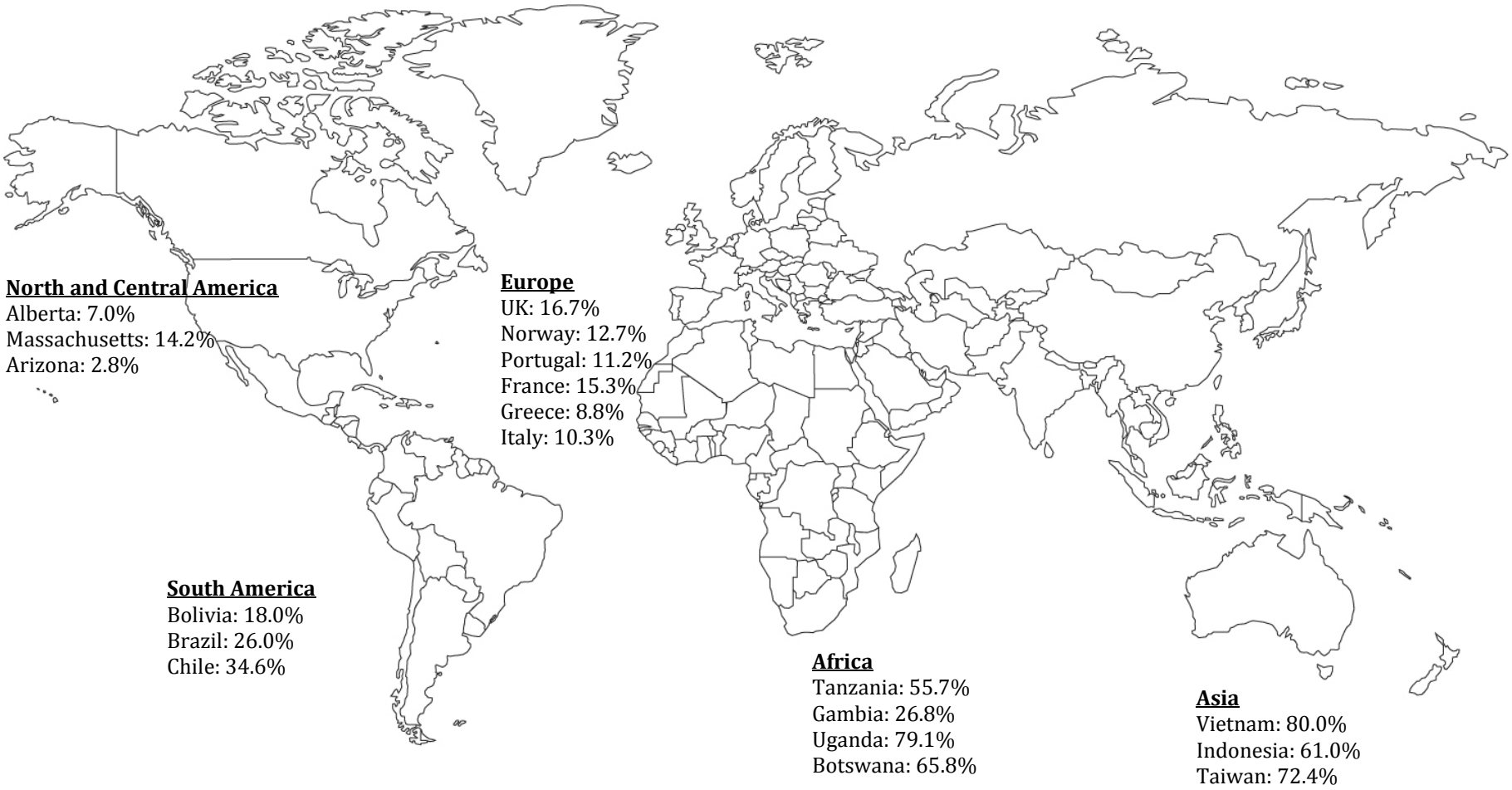


Fig. 1.11. Differences in PCV7 serotype coverage in colonization by geography. Percentages represent proportions of serotypes covered in PCV7 in each country. Proportions are based on selected studies of pneumococcal colonization in children (Huebner *et al.* 1998; Inostroza *et al.* 1998; Joloba *et al.* 2001; Soewignjo *et al.* 2001; Rey *et al.* 2002; Lo *et al.* 2003; Hill *et al.* 2006; Kellner *et al.* 2008; Sa-Leao *et al.* 2009; Cohen *et al.* 2010; Vestrheim *et al.* 2010; Grivea *et al.* 2011; Inverarity *et al.* 2011; Tocheva *et al.* 2011; Moyo *et al.* 2012; Scott *et al.* 2012).

Other environments with high population density such as military barracks, prisons, and detention camps, also have been associated with high carriage rates (Jousimies-Somer, Savolainen & Ylikoski 1989; Hoge *et al.* 1994; Sung *et al.* 1995). Prior hospital admission has also been shown to be a risk factor for pneumococcal colonization, especially antibiotic-resistant pneumococci (Huebner *et al.* 2000b).

1.5.4.2 Socioeconomic Status

Individuals of lower socioeconomic status may be at risk for pneumococcal colonization, possibly due to harsh living conditions, malnutrition, and lack of access to health care facilities. Families living in small areas with lack of health care facilities and lower educational levels may have higher rates and earlier onset of colonization in children (Leach *et al.* 1994; Sung *et al.* 1995). However, some of these studies also compared cohorts of different cultural backgrounds, and other factors such as genetics and ethnicity may affect colonization. St. Sauver *et al.* (St. Sauver *et al.* 2000) did not identify low level of income as a risk factor for colonization, however the study was conducted with the focus of simultaneous pneumococcal colonization (Section 1.6). In a recent study conducted on children in Belgian day care centers, children from families of lower socioeconomic backgrounds showed no significant in pneumococcal colonization rates, but had a higher prevalence of penicillin-nonsusceptible pneumococci belonging to vaccine serotypes and nontypables (Jourdain *et al.* 2011).

1.5.4.3 Seasonality

Colonization rates tend to increase during the winter months for both children and adults (Gray, Converse & Dillon 1980; Darboe *et al.* 2010). A longitudinal study in Israel revealed that in children day care settings, winter months are a risk factor for carriage of single or multi-resistant pneumococci (Yagupsky *et al.* 1998). Different trends of individual serotype prevalence over a year was observed in another one-year longitudinal study in France, where serotype 14 peaked in prevalence in summer while 6B was the most prevalent in the winter (Raymond *et al.* 2000). Rainy months in Africa are also associated with a higher prevalence in carriage (Abdullahi *et al.* 2008). Authors from that work suggested that rainy seasons promote indoor lifestyles, which may enhance horizontal transmission. However, a study in Finland showed a fluctuation in prevalence throughout the year, as opposed to an observable increase in

carriage prevalence in any specific season, as detected by both nasal swabs and aspirates positive for pneumococci (Syrjänen *et al.* 2001).

1.5.4.4 *Virus Infection*

Co-infection with influenza and pneumococcus was associated with increase in pneumococcal adherence (Fainstein, Musher & Cate 1980). Action of viral neuraminidases potentially releases free sialic acid, which has recently been hypothesized to promote biofilm formation in pneumococci. In agreement with this, influenza neuraminidase inhibitors oseltamivir and zanamivir have been shown to reduce pneumococcal colonization in a mouse model (Trappetti *et al.* 2009).

In a longitudinal study in Gambia where sampling was performed once every three months, HIV-positive women are associated with higher rates of pneumococcal colonization determined by pneumococcal-positive cultures (Gill *et al.* 2008b). However, the study could not conclude whether presence of the same serotype in a host in consecutive visits is indicative of repeated acquisition, prolonged duration of carriage, or new acquisition of pneumococcal strain with the same serotypes. One study indicated an association between HIV-positive men and pneumococcal persistence in the nasopharynx, (Rodriguez-Barradas *et al.* 1997). HIV-positive individuals are more likely to carry pneumococci non-susceptible to co-trimoxazole, possibly due to the use of the drug as a prophylaxis in HIV-positive individuals (Rodriguez-Barradas *et al.* 1997; Gill *et al.* 2008a).

1.5.4.5 Smoking

Smoking adults are associated with an increased carriage rate than non-smokers exposed to smoking, which in turn showed a higher carriage rate than individuals not exposed to smoking (Greenberg *et al.* 2006). Children of smoking parents showed higher overall carriage rates and carriage odds ratio than children not exposed to cigarette (Greenberg *et al.* 2006; Cardozo *et al.* 2008). Smoking may also affect viral infection that enhances secondary bacterial colonization (Murphy 2006). Nonetheless, *in vitro* studies demonstrate that pneumococcal adherence was higher in cells of smoking individuals, and exposing non-smoking cells to cigarette-smoke extracts reduced pneumococcal binding (El Ahmer *et al.* 1999).

1.5.5 Colonization as an Immunizing Event

Several studies have shown increases in antibodies specific to capsular and protein antigens such as PspA, PsaA, and pneumolysin following pneumococcal colonization, (Soininen *et al.* 2001; McCool *et al.* 2002; Goldblatt *et al.* 2005; Weinberger *et al.* 2008). Healthy children previously exposed to pneumococci showed an increase in type-specific IgA concentration in saliva, suggesting that carriage induces production of mucosal antibodies (Simell *et al.* 2001). In addition, toddlers in Israel (Weinberger *et al.* 2008) and adults in the United Kingdom (Goldblatt *et al.* 2005) showed increases in levels of IgG specific to carried serotypes such as 14 and 23F, and this rise is associated with a reduction of odds of reacquisition of the same serotype. Through efforts to design a cost-effective vaccine eliciting broad species-wide protection, it was revealed that $CD4^+$ T-cells provide serotype-independent protection from subsequent pneumococcal colonization and rapid clearance from the nasopharynx by macrophages, possibly by the expression of IL-17A (Malley *et al.* 2005; Malley, Srivastava & Lipsitch 2006; Lu *et al.* 2008).

1.5.6 The Role of Capsule in Colonization

During colonization, full encapsulation may hinder exposure of pneumococcal surface adhesins important in host epithelial attachment. In natural populations, nontypable strains with no apparent mechanism for capsule production do colonize (Hathaway *et al.* 2004; Hanage *et al.* 2006). Transparent phase variants with reduced amounts of capsule polysaccharide are associated with a higher capacity for host attachment in an infant rat model, compared to opaque variants with higher capsule expression (Weiser *et al.* 1994). Mutants producing only 20% of capsule had similar concentrations of pneumococci in nasal washes to parental strains (Magee & Yother 2001). Biofilm formation in a non-capsulated strain M11 enhanced biofilm-forming capacity *in vitro*, compared to capsulated derivatives of M11 (Moscoso, García & López 2006). However, strains producing very low amounts of capsule are impaired in biofilm production and enhanced mucosal clearance during colonization (Nelson *et al.* 2007b; Domenech, García & Moscoso 2009). The observation that anti-capsular antibodies reduces carriage rates (Dagan *et al.* 2002), and the differing propensities of varying serotypes to colonize the nasopharynx, suggest that capsule and its chemical structures do play a role in colonization.

1.5.7 Inter-Species Competition and Co-Colonization

Opportunistic pathogens such as *H. influenzae*, *Moraxella catarrhalis*, *Staphylococcus aureus*, and other *Streptococcus* species also colonize the respiratory tract. Children can be co-colonized by these species simultaneously at an early age with high prevalence (Kwambana *et al.* 2011). Interspecies competitions thus are expected to occur for nutrients and attachment space. In addition, interference between species can occur, in which presence of one species protects against the colonization of another. Interaction between species in the same ecological niche is also influenced by the host responses and presence of inhibitory compounds; co-colonization with *H. influenzae* has been shown to enhance pneumococcal phagocytosis by the host immune system (Lysenko *et al.* 2005; Lysenko *et al.* 2007).

Positive and negative correlations between pneumococcal colonization and that of *H. influenzae* have been reported in *in vitro* and population-based studies (Lysenko *et al.* 2005; Pettigrew *et al.* 2008; Margolis, Yates & Levin 2010; Kwambana *et al.* 2011). Pneumococcal H_2O_2 was shown to be toxic for *N. meningitidis*, *M. catarrhalis*, and *H. influenzae* (Pericone *et al.* 2000). Co-existence of *S. pneumoniae* and *H. influenzae* has been shown to promote biofilm formation, however this was a model of middle ear infection (Weimer *et al.* 2010). The main pneumococcal neuraminidase, NanA is able to cleave sialic acid residue from lipopolysaccharides of *N. meningitidis* and *H. influenzae* strains, rendering these colonizers more vulnerable to host complement defences (Shakhnovich, King & Weiser 2002). An inverse relationship between carriage rates of *S. pneumoniae* and *S. aureus* has been observed (Bogaert *et al.* 2004a; Regev-Yochay *et al.* 2004a).

The effect of pneumococcal vaccination may affect replacements of colonization by other species, however thus far observations have been conflicting. With some studies showing an increase in prevalence of *S. aureus* in carriage following a pneumococcal vaccination trial (van Gils *et al.* 2011a), others fail to demonstrate replacements by *H. influenzae* and *M. catarrhalis* (van Gils *et al.* 2011b; Dunne *et al.* 2012).

1.5.8 Intra-Species Competition and Co-Colonization

Competitive interactions between co-colonizing pneumococci have been shown, and may partially explain the emergence of non-vaccine serotypes following PCV7

immunization (Lipsitch *et al.* 2000; Zhang, Auranen & Eichner 2004). Colonization by a resident strain can reduce the colonization of a second strain (Lipsitch *et al.* 2000). However, this inhibition of acquisition is subject to the concentration of the challenge strain. Competitiveness may in part be a serotype-specific property, but other serotype-specific properties such as duration of carriage and transmissibility need to be considered to assess the success of a strain in co-colonization (Melegaro *et al.* 2007). Recently, a longitudinal study in Finland demonstrated that children colonized by a serotype have a 10-fold reduction in acquisition of another serotype, compared to acquisition in un-colonized children (Auranen *et al.* 2010). In contrast, once colonized with 2 serotypes, clearance of a serotype is not dependent on simultaneous carriage of another serotype.

Some mechanisms that may play a role in intra-species competition include release of bacteriocins. Sequence analysis shows sequence variation in the two bacteriocins BlpM and BlpN across serotypes may be grouped into two clusters that may play roles in competition *in vivo* (Dawid, Roche & Weiser 2007). In addition, activation of the pneumococcal competence pathway leads to the upregulation of genes encoding enzymes involved in fratricidal lysis of noncompetent cells (Chapter 5). Hence, differential activation of competence and release of such lytic enzymes may contribute to the lysis of a strain within the nasopharynx by another.

1.5.9 Pneumococcal Biofilm

Biofilms are complex sessile communities comprised of extracellular polysaccharides, proteins, and nucleic acids. Biofilms generally consist of different species co-colonizing as a polymicrobial community. Pneumococci adopt a biofilm mode of growth during colonization and otitis media (Hall-Stoodley *et al.* 2006), in which different transcription and protein expression profiles compared to planktonic pneumococci are reported (Allegrucci *et al.* 2006; Oggioni *et al.* 2006; Sanchez *et al.* 2011). Biofilm formation involves initial attachment, formation of cellular aggregates, and dispersal of single cells to other attachment sites. Mutations causing impaired *in vitro* biofilm formation also impaired colonization in a mouse model (Muñoz-Elías, Marcano & Camilli 2008). There is a positive correlation between the formation of an extracellular matrix and stability of biofilms during murine colonization (Trappetti *et al.* 2011b). Biofilm formation appears to be a general characteristic of pneumococci

from different isolation sites and of different serotypes, although some serotypes (6B, 14, and 19F) have been shown to be more efficient biofilm formers than others (3 and 38). (Tapiainen *et al.* 2010; Camilli, Pantosti & Baldassarri 2011). Drug-susceptible strains also tend to form thicker biofilms.

Biofilm growth provides a favorable environment for adaptation and persistence. Increased antibiotic resistance is a major characteristic of a biofilm, and different mechanisms of this resistance have been proposed, such as reduced diffusion and penetration of drugs through the biofilm matrix and into cells, and the reduced growth rates of cells within a biofilm[19] (Donlan 2000). Isolates from a biofilm have been documented to possess higher MICs to a variety of antibiotics than those in planktonic mode of growth (García-Castillo *et al.* 2007).

The requirement for extracellular DNA for maintaining biofilm architecture and the higher transformation efficiency in biofilm cells, probably mediated by competence activation (Chapter 5), may provide a suitable condition for gene exchange (Li *et al.* 2001; Moscoso, García & López 2006; Oggioni *et al.* 2006). The different environments within a biofilm also contribute to variations in pneumococcal exposure to antibiotics and oxidative stress. These factors may affect genetic transformation and mutation, giving rise to phenotypic variants within a biofilm (Pericone *et al.* 2002; Prudhomme *et al.* 2006; Allegrucci & Sauer 2007; Allegrucci & Sauer 2008).

1.6 Multiple Pneumococcal Colonization

Multiple pneumococcal colonization (or co-colonization) is the simultaneous colonization of pneumococcal strains. Most of the early reports of co-colonization employed culture-based methods and showed differences in its prevalence across continents (Chapter 4). More recent approaches in determining genetic diversity included sensitive DNA-based methods for serotyping (Chapter 3 and 4), as well as strain characterization mainly by PFGE, MLST, and whole-genome sequencing (Chapter 4).

[19] Antibiotics such as β-lactams target metabolically active cells which undergo cell division, as they target proteins necessary for cell wall synthesis. Therefore, metabolically inactive cells would be minimally affected by these drugs.

1.6.1 Effects of Vaccination on Co-Colonization

Two recent studies analyzed the effects of vaccination on co-colonization, one of which showed no significant change in the prevalence of multiple colonization for up to 3 years (Brugger *et al.* 2010), while the other showed a significant reduction in co-colonization rates in the PCV7 era compared to pre-vaccine era (Valente *et al.* 2012). The different observations may be due to the differences in background colonization rates, cohort age and characteristics, as well as laboratory methods of detecting co-colonization. The latter study also demonstrated the effects of herd immunity in reducing rates of co-colonization with vaccine serotypes in non-vaccinated individuals. After PCV7 administration, both studies showed an increase in prevalence of NVT in single or co-colonization. However, due to possibly differences in competitiveness of NVTs, some serotypes were more prominent than others in single colonization than multiple colonizations in the PCV7 era, indicating that composition and relative prevalence of serotypes in co-colonization is not only a reflection of exposure to that serotype. The effects of conjugate vaccines in lowering co-colonization have significant adaptive consequences, as it may reduce the diversity of colonization strains undergoing horizontal transfer of genes encoding advantageous traits (Valente *et al.* 2012).

1.6.2 Horizontal Gene Transfer (HGT)

Pneumococci are naturally competent for genetic exchange (Chapter 5), and multiple pneumococcal colonization facilitates the optimal environment for such exchanges between heterogeneous strains by horizontal gene transfer (HGT). Genes encoding virulence and colonization factors that showed evidence of HGT include *pspA* (Hollingshead, Becker & Briles 2000), IgA1 protease (Lomholt 1995), *pspC* (Iannelli, Oggioni & Pozzi 2002), and *nanA* (Dowson *et al.* 1997; King, Whatmore & Dowson 2005); these are all surface components subjected to selective pressure from the host immune system. In addition, two crucial properties of the pneumococcus, serotype and antibiotic susceptibility, can also be altered by intra- and inter-species HGT of the capsulation locus and antibiotic-resistance determinants, respectively, between pneumococci and other streptococci (Coffey *et al.* 1991; Coffey *et al.* 1998; Brueggemann *et al.* 2007; Croucher *et al.* 2011).

1.6.3 Pneumococcal Supragenome

Comparisons of multiple pneumococcal strains at the genomic level (either by whole-genome sequencing or CGH) reveal the extensive heterogeneity in gene content among these strains (Dopazo *et al.* 2001; Hakenbeck *et al.* 2001; Hoskins *et al.* 2001; Oggioni & Pozzi 2001; Tettelin *et al.* 2001; Brücker *et al.* 2004; Obert *et al.* 2006; Shen *et al.* 2006; Hiller *et al.* 2007; Dagerhamn *et al.* 2008; Donati *et al.* 2010; Croucher *et al.* 2011; Hiller *et al.* 2011). The gene pool of multiple pneumococci therefore exists as a "supragenome" with a size larger than a genome of any single strain (Reviewed in (Ehrlich *et al.* 2005) and (Muzzi & Donati 2011). Using the finite supragenome model, which does not assume an endless introduction of new gene clusters upon the addition of genomes (therefore a non-open genome) (Hogg *et al.* 2007), the pneumococcal supragenome size contains approximately 5,100 clusters, compared to a single genome of ~2,400 clusters (Hiller *et al.* 2007).

The distributed genome hypothesis states that the supragenome consists of a set of core genes found in all strains, in addition to accessory (or distributed) genes found in a subset of strains, and genes unique to a particular strain (Ehrlich *et al.* 2005; Shen *et al.* 2005; Hiller *et al.* 2007). Genome sequence comparison of strains TIGR4 and R6 revealed that some genes and gene clusters are specific to each of the strains, and these unique genes made up 13% of the genome (Brückner *et al.* 2004). Similarly, comparison of TIGR4 and a serotype 3 strain revealed presence of unique gene regions (Oggioni & Pozzi 2001). These clusters exhibit mosaic structures with atypical GC content, suggestive of non-pneumococcal sequences. Similarly, evidence of HGT and microevolution was observed among strains in the nasopharynx of host during a respiratory infection (Hiller *et al.* 2010). Similarities in gene content may not be associated with geographical location and serotype, and strains of serotypes commonly detected in carriage vary in gene contents and genetic background (Hiller *et al.* 2007).

1.6.4 Supragenome in Other Bacterial Species

Intra-species genome diversities are also reported for related streptococcal species such as *S. pyogenes, S. agalactiae, S. mitis, S. thermophilus, S. suis* (Tettelin *et al.* 2005; Brochet *et al.* 2006; Lefébure & Stanhope 2007; Denapaite *et al.* 2010; Donati *et al.* 2010; Johnston *et al.* 2010; Zhang *et al.* 2011). Collectively, it is thought that

the streptococcal supragenome exceeds 6,000 genes, triple to that of a single genome. In addition, other respiratory tract colonizers such as *H. influenzae* (Shen *et al.* 2005; Hogg *et al.* 2007), *S. aureus* (Fitzgerald *et al.* 2001), and *M. catarrhalis* (Davie *et al.* 2011) also show genomic differences across strains.

The extent of diversity varies between species. For example, the supragenome of *Bacillus anthracis* is small; no new genes are introduced after only 4 genomes (Read *et al.* 2002). Similarly, an aphid endosymbiont, *Buchnera aphidicola*, also presents no novel genes after 4 genomes (Mira *et al.* 2010). Such closed pan-genomes may be a consequence of a restricted and isolated niche in which acquisition of foreign DNA material is improbable. In contrast, the opportunistic pathogen *Acinetobacter baumannii* appears to have an open pan-genome associated with its ability to persist in multiple host environments such as lungs, urinary tracts, skin, and plastics (Imperi *et al.* 2011). Similarly, the large streptococcal supragenome is associated with the different host anatomical sites in which they can be isolated and the wide spectrum of diseases they cause in humans and animals. Therefore, differences in supragenome sizes between species may reflect on the species' capacity for gene exchange and transformation, as well as its adaptability in different habitats.

1.6.5 Genetic and Genic Heterogeneity as Means of Adaptation

Distributed genes with known functions encode for transporters, bacteriocins, transcription regulators, antibiotic resistance determinants, efflux proteins, metabolic enzymes, and cell surface proteins (Brückner *et al.* 2004; Obert *et al.* 2006; Hiller *et al.* 2007; Ding *et al.* 2009; Croucher *et al.* 2011; Hiller *et al.* 2011). In addition, other distributed clusters are of hypothetical function, some of which may be beneficial during invasion or colonization (Obert *et al.* 2006; Hiller *et al.* 2007). These variations may enable pneumococcus as a species to increase its metabolic capabilities and adapt in different niches, which may explain its ability to survive and cause infections in multiple host anatomical locations. Genomic comparisons of strains of the same multilocus sequence types (STs) sharing identical housekeeping gene fragments revealed gene content differences closely related clones, and these differences have been shown to be associated with variations in disease potentials in animal models (Silva *et al.* 2006; Hiller *et al.* 2011).

1.6.6 The Significance of Co-Colonization in Adaptation

Simultaneous colonization of pneumococci of different serotypes as well as with other streptococci provides opportunity for intra- and inter-species HGT, enabling the acquisition of beneficial properties (Dowson *et al.* 1990; Dowson *et al.* 1993; Sibold *et al.* 1994). Multiple colonizations can expand the supragenome of a species by introducing accessory, non-core genetic determinants into the population (Shen *et al.* 2006; Hiller *et al.* 2007; Donati *et al.* 2010). This ultimately would confer the population an adaptive advantage for survival in different host environments. Clearly, the simultaneously colonization of different strains is required before such HGT between strains can take place. Therefore, the extent of multiple colonization is a measure of the supragenome and thus the adaptive potential of the organism as a population.

1.7 Work in This Thesis

The works in this thesis include the evaluation of phenotypic and genotypic diversities of pneumococci from different perspectives, including strain diversity in co-colonization events, diversity of the competence-stimulatory peptide, and that of the capsular regulatory gene *cpsB*. The works in this thesis included strains collected from Tanzania as well as the United Kingdom, in addition to reference strains kindly provided by collaborators. The study was conducted by the Centre for Clinical Microbiology of the Royal Free and University College Medical School, University College London.

1.7.1 Chapter Three

Continuous surveillance of prevalent serotypes is crucial in understanding the epidemiology of pneumococcal carriage and infections, which aids in assessing efficacies of current vaccines and designing future serotype-based vaccines. Conventional serotyping methods are costly, time-consuming, and non-specific. Current DNA-based methods require the large number of serotype-specific primer pairs, while microarray studies are costly, limiting its use in selected laboratories. Also, the resolution power of microarrays is also dependent on the serotype-specific targets present in the array, whereas a sequence-based method can easily be modified to identify and characterize a new serotype. Here, we present a novel single-primer PCR and sequence-based method targeting capsular regulatory gene *cpsB* that can correctly identify a high number of serotypes, including those covered in the latest 13-valent conjugate vaccines.

1.7.2 Chapter Four

The phenotypic and genotypic diversities were determined for pneumococci colonizing in a cohort of healthy Tanzanian children under the age of five in Moshi, Tanzania. The selection and systematic characterization of up to 20 colonies per colonization event in this study is greater than most previous studies analyzing pneumococcal diversity. We characterized two crucial phenotypes, serotype and antibiotype. This work also entails genetic characterization by multilocus sequence typing (MLST). We demonstrated the co-colonization of multiple serotypes, some of which also had different antibiotic susceptibilities. MLST revealed the presence of unrelated strains of the same serotype, providing evidence that determining serotype

alone is an underestimate of pneumococcal co-colonization diversity. Through this we have provided a framework and a starting point for performing whole-genome sequencing in estimating the supragenome size of pneumococci within a nasopharynx.

1.7.3 Chapter Five

Pneumococcal competence is mediated by the competence-stimulatorypathway containing a ligand peptide (CSP), which exists in multiple allelic forms. This chapter documents a unique discovery of a high prevalence of CSP4 in Tanzania, a rare pherotype in the developed world. We provide evidence that the differences in proportions of CSP between Tanzania and the United Kingdom is likely to be governed by geography rather than disease potentials of different CSP alleles. We also demonstrated that most co-colonization events containing multiple strains had different CSP alleles, potentially facilitating genetic exchanges between co-colonizing strains.

1.7.4 Chapter Six

It is thought that regulation of capsule expression is crucial to the survival of organism, and some serotypes are more commonly encountered than others. We hypothesize that serotypes differ in their capacity to regulate capsule production, perhaps by the sequence diversity of capsular regulatory genes. With the availability of *cpsB* data from Chapter Three we sought to determine intra-serotype diversity of *cpsB* in a large sample of pneumococci. Irrespective of ST differences, we reveal that serotype 6B contains two divergent clusters based on *cpsB,* compared to serotype 1 that clusters closely together. Comparison of over 300 sequences shows different areas of amino acid conservation and substitutions throughout the length of CpsB, while the majority of amino substitutions within serotype 6B occurred between globular domains, which might affect the structure and folding of the protein.

1.7.5 Chapter Seven

While conducting work described in Chapter Five, we encountered a rare CSP6.1 allele in one of the archived strains from the Royal Free Hospital. Phenotypic and MLST suggest its identity as the newly recognized organism *Streptococcus pseudopneumoniae.* To test the hypothesis of CSP6.1 being a common pseudopneumococcal CSP, presumptive strains of different streptococcal species were

compared in their CSP alleles. We show that CSP6.1 is a common allele in *S. pseudopneumoniae*, and we propose that CSP determination is a potential target in differentiation between these closely related organisms.

CHAPTER TWO: Materials and Methods

2.1 General Overview

Pneumococcal strain collection and archiving is divided into two sections: studies conducted in Tanzania, including nasopharyngeal strain collection and archiving, and studies conducted at the Royal Free Hospital (RFH) in London, United Kingdom, including reference and invasive strain collection, as well as collection of non-pneumococcal reference strains. Strains from Tanzania were included in studies discussed in Chapters 3, 4, 5, and 6, and strains from the UK were included in works discussed in Chapters 3, 5, and 6. Each heading is followed by the location in which the studies were conducted. Preparations for growth media and reagents are described in Section 2.9. Strains included in this thesis are represented in Appendix Table A1. Clinical strains included in the *S. pseudopneumoniae* study are described in the Materials and Methods section in Chapter 7. In addition, non-pneumococcal strains provided by collaborators throughout the work of this thesis are indicated in Table 2.1.

Methods covered in Tanzania were performed by Dr. Ndeky Oriyo, Dr. Sarah Batt, and Dr. John Crump, depending on the cohorts as described below. Methods covered in the Royal Free Hospital, London was performed by Mr. Marcus Leung, unless otherwise described below.

2.2 Pneumococcal Strains from Tanzania

2.2.1 Pneumococcal Isolates from Carriage Studies

Pneumococci isolated from Africa were collected from two separate studies. One was conducted in Moshi, Tanzania, where a longitudinal analysis of multiple pneumococcal carriage in children under the age of 6 was undertaken and headed by Dr. Ndeky Oriyo (Oriyo 2011). A separate study in the Huruma Hospital in Kahe, Tanzania, investigated the effect of azithromycin treatment for trachoma on induction of drug resistance in *S. pneumoniae* (Batt *et al.* 2003). This study was headed and performed by Dr. Sarah Batt. Strains from Moshi were included in works discussed in Chapters 3, 4, 5, and 6. Strains from Kahe were included in works discussed in Chapters 3, 5, and 6. Both studies were conducted by the Department of Medical

Microbiology (now referred to as the Centre for Clinical Microbiology) of the Royal Free Campus UCL Medical School in London. For sections 2.2.3 to 2.2.9, procedures for both studies, where they differ, are described separately.

2.2.2 Pneumococcal Isolates from Invasive Studies

Eighteen invasive strains isolated from Tanzanian children and adults were kindly provided by Dr. John Crump (Crump *et al.* 2011a; Crump *et al.* 2011b). Antibiotic susceptibility data for these strains was provided by Ms. Shanette Nixon of the Kilimanjaro Christian Research Institute. Antibiotic susceptibilities were determined according to CLSI guidelines by Dr. Shanette Nixon of Kilimanjaro Christian Medical College, Moshi, Tanzania. Pneumococci were inoculated onto Dorset's egg medium agar slants (Southern Group Laboratory) and shipped to our London laboratory in ambient temperature (Wasas *et al.* 1998). These strains are included in Materials and Methods section of Chapter 5.

2.2.3 Ethical Approval

Ethical approval for the Tanzania multiple colonization study was initiated, implemented and followed-up by Miss Ndekya Oriyo. Approval was granted by the Tanganyika Planting Company Limited (TPC) Ethical Committee, under application reference number L08E02RB. A copy of the approval letter can be found in Appendix Fig. A1.

Ethical approval for the trachoma study was initiated, implemented and followed-up by Dr. Sarah Batt. Approval was given by the London School of Hygiene and Tropical Medicine (application number 597) and approved by the Ethics Committee of the KCMC, and the Tanzanian National Institute of Medical Research. In addition, consultation was undertaken with Balozi[20] village leaders from the study villages prior to the study. The Balozi leaders were in agreement with all protocols and with their blessing village people were very co-operative. Informed consent was obtained from parents or guardians before samples were taken, in the form of signatures or fingerprints on a consent form printed in English and Kiswahili.

[20] Balozi: Ambassador of a group of 10 houses

2.2.4 Demographic Characteristics

The multiple colonization study was conducted near Moshi, in the Kilimanjaro region of Northeastern Tanzania, between January and December 2003 by Dr. Ndeky Oriyo. The city has a population of 144,739 (2002 census). The study site is a semi-closed community located within a sugar plantation named Tanganyika Plantation Company 25 km from the town of Moshi (Fig. 2.1). At time of study, the plantation has approximately 16,000 hectares of land, with cane cultivation allocated to about 6,100 hectares. The 2,300 employees (majority of which permanent) and their families who lived on the plantation are served by the 60-bed TPC Hospital. The plantation is divided into camps and living accommodation is allocated according to occupation and income. At time of study the employees earned incomes of £28-50 per month. Most of the children in the study were from families with low socioeconomic status in Tanzania, and their houses have on average two bedrooms with shared bathroom and toilet facilities. The average family size is five members. Electricity is provided in residential camps but not in an individual household level. Firewood is used for cooking in greater than 90% of the households. Every camp has a small drug dispensary.

Fig. 2.1 Typical houses within the semi-closed community around the Tanganyika Plantation Company.

The trachoma study was conducted by Dr. Sarah Batt in the Huruma Hospital at the Ibunoki village, in the Rombo district along the foothills of Mount Kilimanjaro in

Northeastern Tanzania between August 2000 and February 2001. Information of the Huruma Hospital is available online (Tanzania Volunteers 2011). The hospital serves mainly residents of the Ibunoki village but also individuals from other regions of Tanzania and neighbouring parts of Kenya. The hospital contained 300 beds and offered free healthcare to individuals.

2.2.5 Recruitment Criteria

Recruitment for the multiple colonization study was organized and supervised by Dr. Ndekya Oriyo. The study cohort consisted of healthy children aged between 1 and 5 years who attended the Mother and Child Health (MCH) Centre at TPC Hospital. These children were eligible for the expanded programme of immunisation (EPI) that does not include a pneumococcal vaccine. The recruitment criteria for the children were that they resided on the plantation, were healthy and had not taken any antibiotics over the 3 months prior to recruitment. A questionnaire was completed by the clinician with information regarding the child's gender, age, medical history and current antibiotic treatment. It also included a continuing record of the antibiotics and other medicines taken within a month of each sampling point. Children that were eligible and for whom, informed consent was obtained were screened for pneumococcal carriage. Children positive for pneumococcal carriage at baseline in January 2003 were swabbed monthly, where possible, for the subsequent 11 months. Thus, children with pneumococcal-positive culture in January/February 2003 were followed longitudinally until December 2003 (Oriyo 2011).

Recruitment criteria for the trachoma study was designed by Dr. Sarah Batt as described (Batt *et al.* 2003). Briefly, a single oral dose of azithromycin was provided to each non-pregnant individual over the age of 5 years and ingestion of the drug was supervised as indicated. (Batt *et al.* 2003).

2.2.6 Specimen Collection

Specimen collection for the multiple colonization study was supervised by Dr. Ndekya Oriyo. Both nasopharyngeal and oropharyngeal swabs were taken to accurately detect rate of pneumococcal carriage and diversity of colonizing serotypes (Masters *et al.* 1958; Capeding *et al.* 1995; Rapola *et al.* 1997; Watt *et al.* 2004; Charalambous, Oriyo & Gillespie 2008). Both nasopharyngeal and oropharyngeal

swabs were taken for the first six months of the study from children by the same laboratory assistant. Due to complaints from families regarding the discomfort of children when receiving nasopharyngeal swabs, only throat swabs were taken for the remaining six months of the study.

Calcium-alginate swabs on a flexible aluminum shaft (Medical Wire and Equipment Limited) and cotton tipped swabs (Technical Service Consultants Limited) were used for nose and throat swabs, respectively. Nose swabs were inserted into the child's nose and sample was collected by sweeping the roof of the nasal cavity. The throat swabs were inserted through the child's mouth and a sample was collected by sweeping the back of his/her tongue. The swab was immediately immersed in approximately 1.0 mL of skim milk-tryptone-glucose-glycerol (STGG) medium (Section 2.9.4) contained in a 7.0 mL plastic screw-topped bijoux bottle. The shaft of the swab was cut with ethanol-flamed, sterilized scissors approximately an inch above the swab so that the screw-top lid could be replaced immediately. Inoculated STGG medium in cool box were taken to the KCMC, where inoculated culture was transferred aseptically without the swab and stored at -70°C.

Single colonies were sub-cultured onto blood agar (purity plate) and pneumococcal colonies from purity plates were archived in STGG medium and sent to the RFH in London, UK. Upon arrival STGG archives were stored at -70°C until the start of this study in October 2008.

In the trachoma study, throat swabs were taken from children under the age of 7 in the communities near the Huruma Hospital by Dr. Sarah Batt. Throat swabs were touched onto the back of the oropharynx and then immediately dipped into 5 mL SGG (STGG without tryptone) medium in bijoux bottles. SGG medium was used instead of the charcoal transport media supplied with the throat swab, bypassing the need for culture before freezing. Two 1 mL aliquots for each specimen were archived at -70 at KCMC in 1.8 mL cryo tubes (Nunc). One aliquot was sent back to the UK and one to be kept in Tanzania.

2.2.7 Isolation of *S. pneumoniae* From Primary Swab

Methods of pneumococcal isolation from primary swabs did not differ between the two carriage studies. Isolation was performed by Dr. Ndeky Oriyo or Dr. Sarah Batt depending on the cohort. Ten microliters of primary swab sample in STGG (or SGG for trachoma study) medium was plated onto a primary blood agar plate containing 5 μg/mL gentamicin (BAG) to select for pneumococcal growth (Sondag, 1977; Dilworth *et al.* 1975; Sondag *et al.* 1977). BAG plates were incubated overnight at 35°C with 5% CO_2.

2.2.8 Identification of *S. pneumoniae*

For multiple colonization study, pneumococci from primary plate was identified by Dr. Ndeky Oriyo based on α-haemolysis, indicated by greenish zone around colony growth, and the characteristic greyish draughtsman-like colonies. Up to 10 presumptive pneumococcal colonies were plated separately on a purity blood agar to confirm pneumococcal growth by addition of optochin disc (Oxoid) and bile solubility testing. Purity plates were incubated in conditions as described above, and a zone of inhibition exceeding 14 mm in diameter (including the disc) was considered optochin-sensitive. Zone sizes of 7-13 mm were optochin intermediate-resistant. A swab of optochin-sensitive pneumococci from each purity plate was inoculated in 1.0 mL STGG and archived in -80°C. Only one colony suggestive of pneumococcal growth per primary plate was selected for the trachoma study.

2.2.9 Sample Numbering

For the multiple colonization study, each child in the cohort was given a number. A strain is numbered by the child number, the site of isolation (NP, nasopharynx, or OP, oropharynx), and the isolate number. For example, 35NP10 represents the 10^{th} pneumococcal colony isolated from the nasopharynx of child 35. Should pneumococci be detected in a children over multiple time points over the duration of sample collection (twelve months, one sampling per month), the isolate number is followed by the first three letters of the month (e.g. 11OP6 Jan and 11OP6 Jun). Strains isolated since July were all from oropharyngeal swabs and hence all were named OP.

For the trachoma study, each child was given a demographic number, the first two digits represents the village, the second two digits the house number, and the digits following hyphenation represents the child starting from the eldest. Hence pneumococcal strain 1003-2 represents one that was isolated from the second eldest child in house 3 of village 10. All samples included in this thesis from the trachoma study were taken at the same time point.

2.3 Pneumococcal Strains from the United Kingdom

Pneumococci isolated from sputum, blood, and CSF collected between the years 1993-2000 at the RFH Microbiology Department were included in Chapters 3, 5, and 6. Pneumococci from sites of infection were isolated and purity plated from qualified laboratory personnel of the Microbiology Department as described below. Pneumococci were archived in STGG after confirmation of pneumococcal isolation by phenotypic properties (Section 2.6).

2.3.1 Isolation of *S. pneumoniae* from Sputum

Isolation of pneumococci from sputum was performed by qualified laboratory personnel of the RFH. Isolates included in the works of this thesis were collected between the years 1993-2010, performed as described in the RFH Standard Operating Procedure (Royal Free Hospital NHS Trust 2011c). Briefly, using a sterile swab, a purulent piece of sputum is selected for dry thin smear preparation. Staining of sample was performed using ProDiff method to visualize and compare proportions of leukocytes (stain violet) and epithelial cells (stain red) in sample for quality assessment. Using low (10X) power objective lens, the number of epithelial cells are calculated, and samples with greater than 25 epithelial cells/lower power field are not analyzed, as the high number of epithelial cells may indicate sample contamination with nasopharyngeal colonizing flora.

For samples with fewer than 25 epithelial cells/lower power field, they are mixed with an equal volume of Sputasol and mixture was vortexed for 30 seconds to homogenize sample. A 10-μL loopful of digested sample was then plated onto a CBA plate and grown overnight as described. Pneumococci were initially identified based on colony morphology and α-haemolysis (Section 2.6.2). A colony was then streaked on a purity plate and an 5-μg optochin disc (Oxoid) was added and grown overnight as described

(Sections 2.2.8 and 2.6.2). Presumptive pneumococci that were optochin-sensitive were tested for bile solubility (Section 2.6.3).

2.3.2 Isolation of *S. pneumoniae* from Blood

Isolation of pneumococci from blood was performed by qualified laboratory personnel of the RFH using the BD BACTEC FX Instrument. Isolates included in the works of this thesis were collected between the years 1993-2000, performed as described in Standard Operating Procedure (Royal Free Hospital NHS Trust 2011a). Briefly, samples are drawn directly from patients into BACTEC culture vials, which are subsequently placed into the instrument and incubated at 35°C. Bacterial growth is determined based on detection of CO_2 in the system. The generated CO_2 reacts with dye in culture vial and affects amount of light absorbed by fluorescent signal in the instrument. When positive growth is indicated, the culture is plated onto CBA overnight and pneumococci identified as described above for sputum samples.

2.3.3 Isolation of *S. pneumoniae* from Cerebrospinal Fluid (CSF)

Isolation of pneumococci from CSP was performed by qualified laboratory personnel of the Royal Free Hospital. Isolates included in the works of this thesis were collected between the years 1993-2000, performed as described in Standard Operating Procedure (Royal Free Hospital NHS Trust 2011b). Briefly, CSF samples are received from the hospital ward in a sterile container and plated onto CBA overnight. Identification of pneumococci is performed as described above for sputum samples.

2.4 Pneumococcal Reference Strains

Twenty-nine pneumococcal strains of known serotypes were received by Prof. Stephen Gillespie from GR MICRO prior to the start of the thesis (Appendix Table A1). Serotypes for these strains were confirmed by the Health Protection Agency (HPA) with the kind assistance of Dr. Karen Broughton, Dr. Siobahn Martin, and Dr. Androulla Efstratiou. Reference pneumococcal strains were kindly provided by Dr. Bruno Pichon (HPA) and Dr. Jutta Loeffler (Centre for Amyloidosis and Acute Phase Proteins, UCL Royal Free Campus). Multilocus sequence typing profiles for HPA strains were provided by the HPA. *S. pneumoniae* ATCC 49619 (serotype 19F) was kindly provided by Mrs. Rajita Dasai (RFH). For all but pneumococci received by Jutta Loeffler, strains were archived in STGG. Pneumococci provided by Dr. Jutta

Loeffler, were archived in Brain Heart Infusion (BHI) medium with 10% glycerol were prepared by Dr. Jutta Loeffler. All strains are stored at -80°C.

2.5 Non-Pneumococcal Reference Strains

Streptococcal strains other than *S. pneumoniae* were kindly provided by Dr. Anna Tymon (Eastman Dental Institute), Dr. Derren Ready of the (Eastman Dental Institute, UCL), Mrs. Rajita Dasai (RFH), and Dr. Clare Ling (HPA). Other bacterial strains were provided by Mrs. Rajita Dasai (RFH) and Ms. Sarah Thurston (Respiratory Medicine, UCL) (Table 2.1). Frozen archives in STGG were prepared as described above, except for *H. influenzae*, where chocolate blood agar (Oxoid) was used instead of CBA for overnight incubation[21]. Strains provided by Dr. Derren Ready were archived from CBA in frozen medium containing 50% BHI, 30% Phosphate Buffered Saline, pH 7.2 (Gibco), and 20% glycerol (Sigma).

Table 2.1. Non-pneumococcal strains provided for works included in this thesis

Species	Type Strain	Provided by
Streptococcus anginosus	NCTC 10713	Derren Ready (UCL)
Streptococcus gordonii	NCTC 7865	Derren Ready (UCL)
Streptococcus intermedius	NCTC 2227	Derren Ready (UCL)
Streptococcus mitis	NCTC 12261	Derren Ready (UCL)
Streptococcus mutans	NCTC 10449	Derren Ready (UCL)
Streptococcus parasanguinis	NCTC 55898	Derren Ready (UCL)
Streptococcus salivarius	NCTC 8618	Derren Ready (UCL)
Streptococcus sanguinis	NCTC 7863	Derren Ready (UCL)
Streptococcus sobrinus	NCTC 12279	Derren Ready (UCL)
Streptococcus cristatus	ATCC 51100	Derren Ready (UCL)
Streptococcus agalactiae	ATCC 12401	Rajita Dasai (NHS)
Streptococcus pseudopneumoniae	ATCC BAA-960	Clare Ling (HPA)
Streptococcus oralis	NCTC 11427	Anna Tymon (UCL)
Haemophilus influenzae	ATCC 10211	Rajita Dasai (NHS)
Staphylococcus aureus	ATCC 25923	Rajita Dasai (NHS)
Klebsiella pneumoniae	ATCC 13883	Rajita Dasai (NHS)
Moraxella catarrhalis	ATCC 25238	Sarah Thurston (UCL)

2.6 Identification of *S. pneumoniae*

From frozen culture archived in -80°C in freezer in RFH, confirmation of pneumococcal identity by α-haemolysis, optochin sensitivity and bile solubility were performed. A proportion of pneumococci were also tested for agglutination with

[21] CBA plates are unsatisfactory for growth of *H. influenzae*, and chocolate blood agar with already-lysed erythrocytes is more appropriate.

anticapsular antibodies. The pneumococcus is generally α-haemolytic, optochin sensitive, bile-soluble. Agglutination with anticapsular antibodies depends on the presence and type of the capsule. These tests are described below.

2.6.1 α-Haemolysis

Presumptive pneumococci are indicated by α-haemolysis, seen as a greenish zone around colony growth. This can be differentiated from β-haemolysis, where the clearing of the agar colour can be seen, and γ-haemolysis, where no colour change is observed (Fig. 2.2). α-hemolysis characteristic is not a definitive property of the pneumococcus, as it is seen in other streptococcal species[22] (Rouff & Bisno 2010).

Fig. 2.2. Haemolytic characteristics of Streptococci. *Streptococcus pneumoniae* is α-haemolytic (left), which can be differentiated with β-haemolysis (center), and γ-haemolysis (right), by their effect on red blood cells on blood agar.

2.6.2 Optochin Susceptibility

For each strain, a 5-μg optochin disc (Oxoid) was placed on purity plate. Optochin-susceptibility was not tested directly from frozen STGG culture, as glycerol may induce optochin-resistance (Robson *et al.* 2007). Optochin susceptibility is determined by the zone of inhibition around the disc (O'Brien *et al.* 2003). After overnight incubation, the strain is considered optochin-susceptible if a zone of inhibition exceeds 14 mm, optochin-nonsusceptible if zone of inhibition is 7-14 mm, and optochin-resistant if the zone of inhibition is ≤ 7 mm. All but one strain (strain N452) included in this study was optochin sensitive based on zone size. Subsequent

[22] α-haemolytic streptococci include but not limited to *S. mitis*, *S. oralis*, *S. sanguinis*, and *S. gordonii* (Rouff & Bisno 2010).

phenotypic and genetic characterization suggests N452 belongs to *S. pseudopneumoniae* (Chapter 7).

2.6.3 Bile Solubility

Two variations of the bile solubility tests, the colony and suspension procedures were performed, depending on colony sizes. The colony procedure is preferable when pneumococcal colonies are large or mucoid. The suspension procedure is recommended when pneumococcal colonies are small (Health Protection Agency 2011). Both procedures were performed as described below. Reference *S. pneumoniae* strain ATCC 49619, instead of NCTC 12977 as recommended by the HPA, was tested in parallel as a control.

2.6.3.1 Colony Bile Test

A loopful of 2% sodium deoxycholate (Sigma-Aldrich) was placed onto a large or mucoid colony on purity plate, and plate was incubated at 35°C with 5% CO_2 for 15-30 mins. A flattening of colony indicates that strains of that colony is bile-soluble, and is suggestive of pneumococci. Intact colonies after incubation suggest that strains are bile-insoluble.

2.6.3.2 Broth Bile Test

From purity plate, a loopful of colonies was suspended into 500 μL of 0.9% saline (Oxoid) to make a thick suspension. Five hundred microliters of 10% sodium deoxycholate (Sigma) was added to the test samples, while 500 μL of saline was added to a negative control. Tubes were then placed in 35°C with 5% CO_2 for 15-30 mins. A clearing of the suspension following incubation is indicative of bile-solubility.

2.6.4 Serotyping Using Anti-Capsular Sera

Pooled (Pneumotest Kit) and type-specific antisera (Statens Serum Institut, Copenhagen) were used for serotyping. Pneumokit includes antisera specific for serotypes or serogroups included in the 23-valent polysaccharide vaccine.

Serotyping performed was based on the chessboard method (Sørensen 1993). The kit contains 12 pooled antisera (A-F, and H, and P-T), each containing sera specific to a

group of serotypes or serogroups. The kit is designed such that a vaccine serotype/group will react with one pool from A-F and H and one from P-T. Suspension of each sample was to be mixed with each pool individually and sequentially, first with pools A-F and H, followed by pools P-T. Pools containing serotypes/groups most prevalent in carriage and disease are tested first. Thus, the order is A→B→C→H→D→E→F, then P→Q→R→S→T. In brief, a loopful of multiple pneumococcal pure colonies was re-suspended in 50 µL of PBS, pH 7.2 (Gibco). Two µL of the suspension was added to a plain glass slide, and 1 µL of pooled antiserum was added adjacent to the drop of suspension. A coverslip was then used to mix the two droplets and applied onto the slide, and the slide was then viewed under light microscope.

A positive result is indicated by the agglutination of pneumococci under the light microscope when mixed with sera of a particular pool. The serotype/group was recorded as the type/group reactive to a combination of 2 pools (Table 2.2). For example, a sample in which agglutination was found when exposed to pool B and Q belongs to serogroup 6. Pneumococci belonging to a serogroup where multiple serotypes are present, type-specific antisera were used to identify the serotype of these isolates.

Table 2.2. Serotypes/groups reactive to antiserum pools included in the Pneumotest Kit

Pool	A	B	C	D	E	F	H
P	1	19	7				14
Q	18	6					23
R	4	3		9	12		
S	5	8			10	17	15
T	2		20	11	33	22	
Neg[a]			Group 1	Group 2	Group 3	Group 4	Group 5

[a] Positive reactions to only one pool of A-F and H are in one of these groups:
Group 1: Serotypes/groups 24, 31, 40
Group 2: Serotypes/groups 16, 36, 37
Group 3: Serotypes 21, 39
Group 4: Serotypes/groups 27, 32, 41
Group 5: Serotypes/groups 13, 28

2.7 Antibiotic Susceptibility Tests

2.7.1 Disc Diffusion Test

The Kirby-Bauer disc diffusion method was performed by Dr. Ndekya Oriyo in Tanzania, and was performed to determine susceptibilities to penicillin, amoxicillin, co-trimoxazole (25 μg), tetracycline (30 μg), chloramphenicol (30 μg), erythromycin (15 μg), and amoxicillin, according to guidelines set out by the Clinical and Laboratory Standards Institute (CLSI) (Clinical 2008). An oxacillin disc (1 μg) was used to assess penicillin and amoxicillin susceptibilities. To perform disc testing, colonies from an overnight-grown CBA plate were picked with a cotton swab and inoculated into 0.9% saline to a density equivalent to turbidity of 0.5 McFarland standard by comparing with a control 0.5 McFarland standard tube (bioMerieux). A cotton swab was immersed into the suspension was then streaked to cover the entire area of MHB agar plates. Following streaking antibiotic discs were placed onto MHB and plates were incubated as described above. Table 2.3 shows the susceptible, intermediate-resistant, and resistant breakdown for each antimicrobial tested.

Table 2.3 Zone diameter interpretive standards for *S. pneumoniae* according to Clinical Laboratory and Standards Institute guidelines

Antimicrobial	Sensitive (mm)	Intermediate (mm)	Resistant (mm)
Oxacillin (1 μg)	≥ 20	-	-
Co-trimoxazole (25 μg)	≥ 19	16-18	≤ 15
Tetracycline (30 μg)	≥ 23	19-22	≤ 18
Chloramphenicol (30 μg)	≥ 21	-	≤ 20
Erythromycin (15 μg)	≥ 21	16-20	≤ 15

2.7.2 Minimum Inhibitory Concentration Determination

E-test was performed by Dr. Ndekya Oriyo in Tanzania and was confirmed by Marcus Leung in RFH. For strains non-susceptible to penicillin and co-trimoxazole based on disc diffusion, E-test performed with benzylpenicillin and co-trimoxazole strips (AB bioMerieux), following CLSI procedures (Clinical and Laboratory Standards Institute 2008). *S. pneumoniae* ATCC 49619[23] was included as a control strain. An overnight pure culture from CBA was swept and inoculated into 0.9% saline (Oxoid) to 0.5 McFarland standard. This turbidity was visually compared with

[23] ATCC 49619 with penicillin and co-trimoxazole MIC of 0.25-1 μg/mL and 0.125-1 μg/mL, respectively (Clinical and Laboratory Standards Institute 2008).

0.5 MacFarlane control standard (bioMerieux). The inoculated suspension was then plated onto MHB agar plate (Oxoid). One E-test strip was placed onto the center of each plate. For each strain triplicates were performed (three for penicillin, three for co-trimoxazole per strain). MHB containing E-test strips were incubated overnight at 35°C with 5% CO_2. MIC was read on the confluent grown plate the next day, measured by the concentration scale of the strip where growth inhibition meets (Fig. 2.3). In cases where MICs differ within the triplicate by two-fold, the median MIC values were indicated in results. For greater than two-fold discrepancies, triplicate tests were re-performed. Table 2.4 shows the susceptible, intermediate-resistant, and resistant breakpoints for penicillin and co-trimoxazole as interpreted by CLSI guidelines (Clinical and Laboratory Standards Institute 2008).

Table 2.4 Minimum Inhibitory Concentration (MIC) interpretive standards for *S. pneumoniae* according to CLSI guidelines

	MIC breakpoints (µg/mL)		
Antimicrobial	**Sensitive**	**Intermediate**	**Resistant**
Penicillin (non-meningitis)	≤ 0.06	0.12-1.0	≥ 2.0
Co-trimoxazole	≤ 0.5	1.0-2.0	≥ 4.0

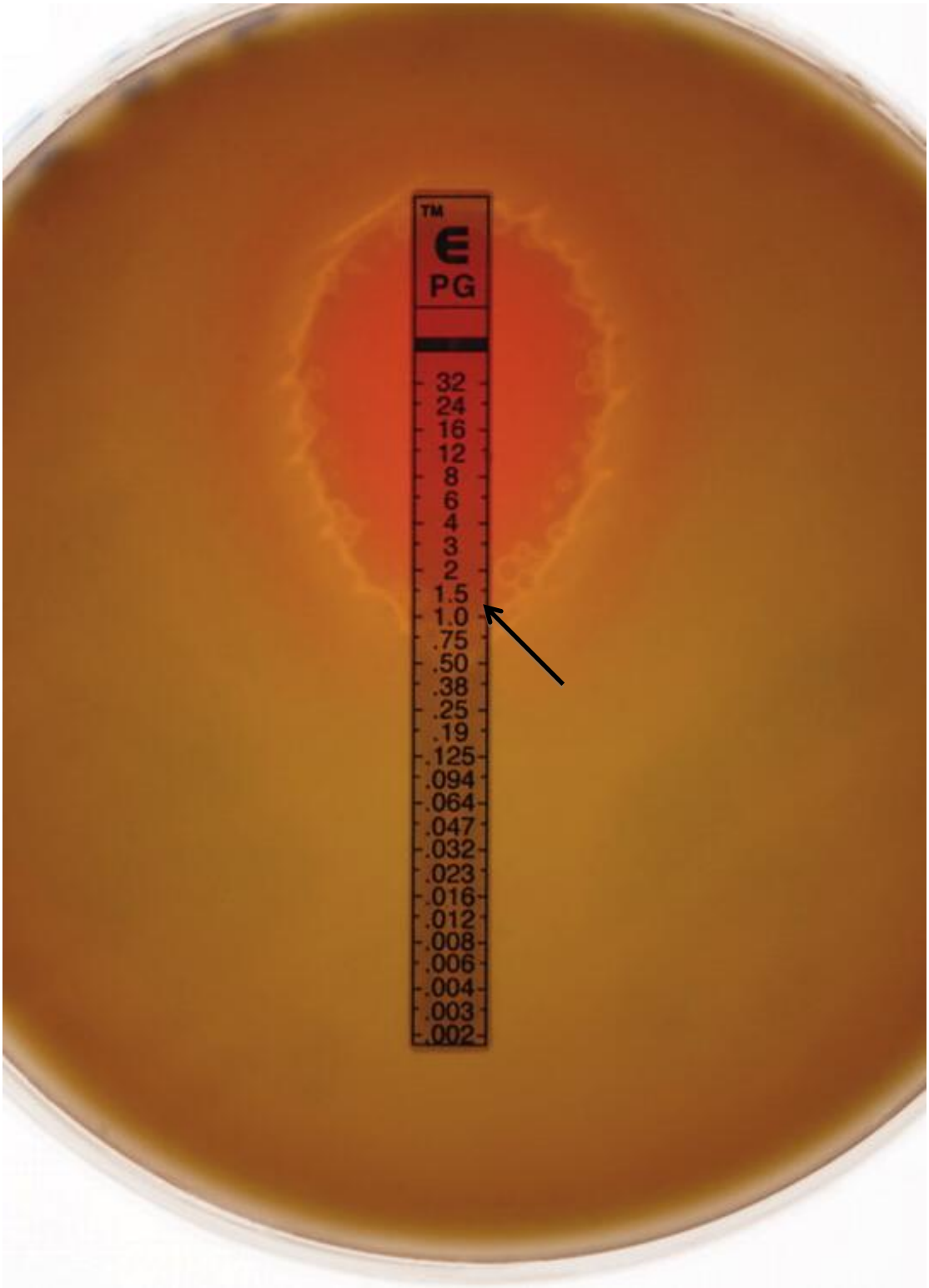


Fig. 2.3 MIC determination using E-test strips on a confluent growth of pneumococci. MIC of a strain is indicated as the concentration scale of the strip where growth inhibition meets. The strip shown is the benzylpenicillin E-test strip. Arrow indicates the MIC value for this strain. Figure modified from Morand and Muhlemann (Morand & Mühlemann 2007). Proceedings of the National Academy of Science of the USA allows the use of original figures including one shown above for noncommercial and educational use without requesting permission.

2.8 Genetic Analysis of Pneumococcal Strains

The works in this thesis include sequence analysis of capsulation regulatory gene *cpsB* (Chapter 3), housekeeping genes in multilocus sequence typing (Chapters 4 and 7), gene encoding the competence pathway ligand peptide *comC* (Chapter 5 and 6), and PCR amplification of 16S ribosomal gene from non-pneumococcal organisms (Chapter 3). The sections that follow describe the procedures required to extract DNA from intact cells, followed by PCR amplification, and sequence analysis. The steps described below are standard methodologies applicable to all sequence analysis across all gene regions analyzed. For specific procedures, please refer to the materials and methods sections in respective chapters.

2.8.1 Standardized DNA Extraction by Heat Lysis Method

Overnight pneumococcal colonies from purity plates were harvested with a sterile loop (Nunc) and inoculated into 50 µL 1X PBS, pH 7.2 (Gibco). Microcentrifuge tubes were heated at 95ºC in Accublock Digital Dry Bath (Labnet) for 5 mins to induce cell lysis. Suspension was then centrifuged at 10,000 x g for 5 mins. A 1/10 dilution was made by mixing 10 µL supernatant containing genomic DNA with 90 µL PBS (Gibco) or PCR-grade water (Invitrogen). Diluted genomic DNA was stored in 4ºC for up to six months.

2.8.2 DNA Quantification and Quality Verification

Purified DNA was quantified using the NanoDrop 2000 Spectrophotometer (Thermo Scientific). Briefly, the instrument was first calibrated with either 2 µL of PBS or PCR-grade water depending on the elutant of the DNA sample to be assessed. The sample loading dock was then cleaned with a tissue wipe and the same volume of DNA sample was examined for its concentration and purity as measured by a 260/280-nm absorbance ratio. A ratio with a range of 1.8-2.0 is considered to be pure.

2.8.3 Standardized Polymerase Chain Reaction (PCR)

Each PCR reaction contained 50 µL master mix consisting of 5 µL 10X PCR buffer (Invitrogen), five µL $MgCl_2$ (Invitrogen), three µL dNTP (Promega), two µL of each oligonucleotide primer (Sigma-Aldrich), and 0.3 µL *Taq* polymerase (Invitrogen), and 32.7 µL DNase/RNase-free distilled water (Invitrogen). Stock and master mix concentrations of PCR components are presented in Table 2.5.

Table 2.5. PCR components and stock and master mix concentrations

Component (Stock concentration)	Volume in Master Mix (µL)	Master Mix Concentration (in 50 µL)
PCR Buffer (10X)	5	1 X
$MgCl_2$ (50 mM)	3	3 mM
dNTP (5 mM)	3	0.3 mM
Primers (10 µM)	2	0.4 µM each
Taq polymerase (5 U/µL)	0.3	0.03 U/µL

Table 2.6 lists primers used for conventional PCR performed in this thesis and the amplification extension time for each primer pair. Other than differences between Tm and extension time, all PCRs performed had an initial denaturation step at 94 ºC, followed by 30 cycles of amplification. Amplification included 30 sec. of denaturation at 94 ºC, thirty sec. of annealing, and extension at 72 ºC. A final 4-minute extension step at 72 ºC is included. For works pertaining to quantitative PCR (qPCR) amplification in works of Chapter 7, sequences of primers and quantitative PCR (qPCR) probes are included in Section 7.2.

Table 2.6 Conventional PCR primers used in works of thesis

Primer	Chapter	Sequence (5' – 3')	Tm(°C)	Size (bp)	ET (Sec)[a]	Reference
aroE-fw	3, 5	GCC TTT GAG GCG ACA GC	50	405	45	(Enright & Spratt 1998)
aroE-rv	3, 5	TGC AGT TCA (G/A)AA ACA T(A/T)T TCT AA	50	405	45	(Enright & Spratt 1998)
gdh-fw	3, 5	ATG GAC AAA CCA GC(G/A/T/C) AG(C/T) TT	50	460	45	(Enright & Spratt 1998)
gdh-rv	3, 5	GCT TGA GGT CCC AT(G/A) CT(G/A/T/C) CC	50	460	45	(Enright & Spratt 1998)
gki-fw	3, 5	GGC ATT GGA ATG GGA TCA CC	50	483	45	(Enright & Spratt 1998)
gki-rv	3, 5	TCT CCC GCA GCT GAC AC	50	483	45	(Enright & Spratt 1998)
recP-fw	3, 5	GCC AAC TCA GGT CAT CCA GG	50	450	45	(Enright & Spratt 1998)
recP-rv	3, 5	TGC AAC CGT AGC ATT GTA AC	50	450	45	(Enright & Spratt 1998)
spi-fw	3, 5	TTA TTC CTC CTG ATT CTG TC	50	474	45	(Enright & Spratt 1998)
spi-rv	3, 5	GTG ATT GGC CAG AAG CGG AA	50	474	45	(Enright & Spratt 1998)
xpt-fw	3, 5	TTA TTA GAA GAG CGC ATC CT	50	486	45	(Enright & Spratt 1998)
xpt-rv	3, 5	AGA TCT GCC TCC TTA AAT AC	50	486	45	(Enright & Spratt 1998)
ddl-fw	3, 5	TGC (C/T)CA AGT TCC TTA TGT GG	50	441	45	(Enright & Spratt 1998)
ddl-rv	3, 5	CAC TGG TG(G/A) AAA CC(A/T) GGC AT	50	441	45	(Enright & Spratt 1998)
cpsA-fw	3	GCA GTA CAG CAG TTT GTT GGA CTG ACC	54	160	30	(Pai, Gertz & Beall 2006)
cpsA-rv	3	GAA TAT TTT CAT TAT CAG TCC CAG TC	54	160	30	(Pai, Gertz & Beall 2006)
cpsB-fw	3, 6	GCA ATG CCA GAC AGT AAC CTC TAT	65	1061	90	This work
cpsB-rv	3, 6	CCT GCC TGC AAG TCT TGA TT	65	1061	90	This work
cps13wzy-fw	3, 6	GAT GGG AAA ATA CGA TAT GCT C	61	309	50	Kong (2005)
cps13wzy-fw	3, 6	AAC TCC ATG ACA AAA CTC CAG C	61	309	50	Kong (2005)
cps34I6bh-fw	3, 6	GCT TTT GTA AGA GGA GAT TAT TTT CAC CCA AC	54	408	45	(Pai, Gertz & Beall 2006)
cps34I6bh-rv	3, 6	CAA TCC GAC TAA GTC TTC AGT AAA AAA CTT TAC	54	408	45	(Pai, Gertz & Beall 2006)
cps6cwciN-fw	3, 6	TAC CAT GCA GGG TGG AAT GT	55	1800-2000	150	(Park *et al.* 2007)
cps6cwciN-rv	3, 6	CCA TCC TTC GAG TAT TGC	55	1800-2000	150	(Park *et al.* 2007)
comC-fw	4, 5	GAT AAA ATT CTC TCA ACT GT	54	~300	30	This work
comC-rv	4, 5	GGT AAC TGT GAC TAA TAA TT	54	~300	30	This work
16slf-fw	6	AGT TTG ATC CTG GGC TCA G	55	~1300	120	(Woo *et al.* 2001)
16slf-rv	6	AGG CCC GGG AAC GTA TTC AC	55	~1300	120	(Woo *et al.* 2001)

[a] ET: Extension time

2.8.4 Agarose Gel Electrophoresis

Electrophoresis gel with 1.5% agarose was made by adding 0.45 g of agarose (Bioline) to 30 mL 1 x TBE buffer. Mixture was heated in a microwave oven (Sharp Electronics) to boiling to melt agarose powder. Ethidium bromide (Fluka) was added to a final concentration of 0.05 µg/mL. Each lane contains 3 µL containing PCR reaction and loading buffer (Qiagen). Gel bands were visualized in either ImageQuant 300 (GE Healthcare Life Sciences) or G:BOX (SynGene) gel imaging systems. Hyperladder I (Bioline), 100-bp (Invitrogen) or 1000-bp ladders (Invitrogen) were used for band size standardizations for all electrophoresis runs depending on reagent availability.

2.8.5 PCR Amplicon Purification

Purification of PCR products was performed using QIAquick PCR Purification Kit (Qiagen), according to manufacturer's instructions. The elutant containing purified DNA was stored in -20ºC.

2.8.6 Cycle Sequencing

The same primers used for PCR amplification were employed for cycle sequencing. Each sequence reaction contains 1.8 µL 5X Big Dye Sequencing Buffer (Applied Biosystems), two µL of forward or reverse primer, and 0.5 µL of Big Dye Sequence Terminator v3.1 (Applied Biosystems). Sterile distilled water (Invitrogen) was added to a final volume of 7.5 µL of master mix, followed by the addition of 2.5 µL of purified PCR amplicon. Sequencing reactions were added into 96-well plate (Applied Biosystems) and were subject to an initial 96 ºC step, followed by 25 cycles of 96 ºC for 30 sec., 50 ºC for five sec., and 60 ºC for four mins. Each sample is run in quadruplicate: duplicate in the forward direction and duplicate in the reverse direction. Completed sequencing reactions were kept at 4ºC until ethanol precipitation (see below). Ethanol precipitation was performed no later than 3 days after cycle sequencing to ensure fluorescent dye from Big Dye does not dissipate over time.

2.8.7 Ethanol Precipitation

Ethanol precipitation was performed directly on 96-well plates (Applied Biosystems) to cycle-sequenced product to desalt and concentrate amplified single-stranded products. Each sequencing reaction (10 µL) was mixed with 90 µL of precipitation

solution containing 62.5 µL 95% ethanol (Sigma-Aldrich) with 24.5 µL HPLC water (VWR) and 3 µL of sodium acetate, 3M, pH 4.6 (Applied Biochem). Precipitation reactions in 96-well plates were centrifuged at 4,000 x g for 45 mins at 4°C in the ALC PK120 centrifuge (ALC). The solution was discarded by gently tapping the 96-well plate upside down to not dislodge invisible DNA pellet. Excess salt in precipitation was removed by the addition of 200 µL of 70% ethanol, followed by centrifugation of 4,000 x g for 15 mins at 4°C. Ethanol was discarded in the same fashion. To remove excess ethanol, the 96-well plate was spun upside-down at 800 x g for 1 min at 4°C to remove excess ethanol without removing DNA pellet, and put on 95°C hot plate for 10 seconds. Fifteen µL of HiDi formamide [24] (Applied Biosystems) was added to each reaction well. Plates containing purified DNA pellets were either immediately analyzed on the sequence analyzer (see below) or stored at -20°C until sequence analysis. Stored plates were processed no later than a week to ensure fluorescence signal is sufficient for sequence analysis.

2.8.8 Sequence Analysis

Sequencing plates were loaded onto Sequence Analyzer 3130 (Applied Biosystems) with the 50 cm, four-capillary array set up. Polymer POP-7 (Applied Biosystems) was used for sequencing reactions over the course of this thesis. Each set of 4 reactions took approximately 60 minutes. Therefore, a full 96-well plate containing quadruplicates of 24 samples would take approximately one day to complete. Raw data files were automatically converted into .ab1 files that can be analyzed by sequence analysis software Bionumerics.

2.8.9 Bionumerics

Bionumerics software versions 3.0 or 6.5 (Applied Maths) were used to perform sequence analysis, sequence trimming, and sequence data storage. Ab1 sequence files generated by sequence analyzer were loaded onto Bionumerics interface and multiple alignment (from each of four cycle sequencing reactions per sample) was performed to generate a consensus sequence. Alignment was analyzed over the length of the consensus. Should undefined or overlapping peaks appear within the fragments

[24] Addition of formamide improves sequencing reactions by weakening hydrogen bonding between nucleotides (Zhang, Hu & Deisseroth 1991).

required for analysis (with MLST loci, *cpsB,* or *comC*), sequences were classified as of poor quality, and cycle sequencing was repeated from purified PCR products. If necessary, plating from frozen culture and PCR were also repeated. During the course of this thesis Bionumerics was upgraded to version 6.5. Database was transferred from the old version into the new.

2.9 Growth Culture and Media

2.9.1 Columbia Blood Agar (CBA)

Blood agar with gentamicin was prepared by Dr. Ndekya Oriyo in Tanzania. CBA plates used in London do not contain gentamicin as archived samples were selected for pneumococcal growth in Tanzania; CBA plates used in RFH were purchased from the manufacturer (Oxoid), and were stored in 4ºC until use. All plates were inoculated prior to the expiration date indicated on package.

2.9.2 Mueller Hinton Agar with Horse Blood (MHB)

MHB plates used in Tanzania was prepared by Dr. Ndekya Oriyo by mixing 38 g of Muller-Hinton powder (Oxoid) to 950 mL of distilled water and autoclaved at 121°C. MHB plates used in RFH were prepared by the manufacturer (Oxoid), and were stored in 4ºC until use. All plates were used prior to the expiration date indicated on package.

2.9.3 Brain Heart Infusion (BHI) Broth

BHI broth was made by dissolving 37 g of BHI (Cosmos Biomedical) in distilled water. Broth was autoclaved at 121°C at 15 p.s.i. for 15 minutes.

2.9.4 Skim Milk-Tryptone-Glucose-Glycerol (STGG) Medium

2.9.4.1 Rationale for its Use

The requirement for detailed analysis of nasopharyngeal sample calls for the optimal recovery of pneumococci following long-term storage. Sampling of such clinical samples in the developing world should also provide a transport medium which can maintain the viability of the samples (Leach *et al.* 1997). STGG (skim milk, tryptone, glucose, glycerol) medium was designed by Gibson *et al.* (Gibson & Khoury 1986) as a transport medium. This study showed that 97 bacterial species, including *S.*

pneumoniae, stored in STGG was viable for at least 40 months. Similarly, cell count was maintained in this medium for at least 3 years with repeated freeze-thaw cycles (Kaijalainen *et al.* 2004). There was no detectable difference in recoveries of pneumococci between direct plating and STGG storage at both -70°C and -20°C for at least 9 weeks; however long-term storage is preferable at -70°C (O'Brien *et al.* 2001). A recent study by Hare *et al.* (Hare, Smith-Vaughan & Leach 2011) showed that pneumococci in STGG remain viable for up to 12 years. The correlation between direct plating and STGG in detecting viable pneumococci even in low cell concentration is important for characterization of pneumococci in low abundance (Gratten *et al.* 1994; Brugger *et al.* 2010). Put together, the transport and long-term storage of pneumococci in STGG provides optimal and representative recovery of nasopharyngeal sampling, bypassing the requirement for immediate pneumococcal identification and characterization following sampling. Thus, a more extensive analysis can be performed at a later time without significant loss of viability.

2.9.4.2 STGG Components

Each component of STGG contributes to the preservation of cells. While skim milk is included in STGG to replace glass beads for bacterial agglutination, skim milk alone is capable of sustaining viable pneumococci for at least 16 months (Siberry *et al.* 2001). Tryptone soya broth is a nutrient source when cells are grown in this medium. The inclusion of tryptone to the original SGG (skim milk, glucose and glycerol) medium has been demonstrated to extend preservation of pneumococci at higher storage temperatures (Charalambous *et al.* 2003). Glucose provides a protective effect to cells from freezing (Gherna 1981), and glycerol is capable of diffusing through the cell membrane and also provides cellular protection from freezing. The properties of STGG enable a slow cooling rate of 1.5°C/min when stored in -70°C, which allows optimal survival and recovery of bacterial organisms during and after long-term freezing (Gherna 1981; Gibson & Khoury 1986).

2.9.4.3 STGG Preparation

Three grams of tryptone soya broth (Oxoid) two grams of skim milk powder (Sigma-Aldrich), and 0.5 g glucose (Sigma-Aldrich) were added into a 250 mL Pyrex bottle (Duran). Ten milliliters of glycerol (Sigma) and 100 mL distilled water was added. The mixture was heated to 70°C and stirred with a magnetic stirrer until all

components are dissolved. Aliquots containing 1 mL were dispensed into 1.8 mL cryo-tubes (Nunc). Tubes were either autoclaved at 15 p.s.i. at 121°C for 10 mins, or tyndallized (see section 2.7.4.4). The sterilized medium was stored in 4°C until use. Sterile STGG was used within six months of preparation to ensure stability of inoculant after sterilization (O'Brien *et al.* 2001).

2.9.4.4 Tyndallization

Tyndallization is a process for sterilizing under atmospheric pressure as an alternative to autoclaving if such a pressurized sterilization device is not available or the medium would be destroyed by high temperature. At times where autoclaving facilities are out of service, tyndallization was employed. Utilizing the germination process, bacterial spores are killed under repeated boiling. STGG was boiled for 15 minutes, then incubated in 37°C with 5% CO_2 overnight. The remaining spores that survived the boiling process would germinate and become vegetative cells during the overnight incubation. This boiling and incubating process was repeated twice to maximize elimination of spore-forming contaminants. After tyndallization STGG was stored in 4°C until use. Immediately prior to the need for long-term storage of bacterial cells, a sample of tyndallized STGG was plated onto Columbia blood agar and incubated overnight to confirm the absence of contaminants.

2.10 Preparation of Solutions, Buffers, and Other Reagents

2.10.1 Optochin Discs

Optochin discs (Oxoid) were stored at 4°C until use.

2.10.2 Bile (Sodium Deoxycholate)

For the colony solubility method, two percent bile solution was prepared by dissolving 0.2 g of sodium deoxycholate (Sigma-Aldrich) into 10 mL sterile distilled water. For the broth solubility method, ten percent bile solution was prepared by dissolving 1 g of sodium deoxycholate into 10 mL sterile distilled water. Solution was stored in 4°C in the dark to keep away from light source.

2.10.3 Working Primer Stocks

Primers were received from manufacturer in powdered forms (Sigma-Aldrich). Stock solution of primers were made with DNase/RNase-free distilled water (Invitrogen) to

a concentration of 100 µM and stored at -20°C. Working stocks were made by mixing 20 µL of 100 µM primer stocks with 180 µL DNase/RNase-free distilled water (Invitrogen) to a concentration of 10 µM. Working stocks were stored at 4°C.

2.10.4 dNTP Working Stocks

Stock dNTPs were supplied by Promega. Separate tubes of dATP, dGTP, dCTP, and dTTP were supplied at a concentration of 100 mM. Working stocks were made by mixing 50 µL of each dNTP with 800 µL DNase/RNase-free distilled water (Invitrogen) to a 1000 µL working stock 5 mM dNTP mix.

2.10.5 Tris-Borate EDTA (TBE) Buffer

TBE buffer (0.5 X) was prepared by adding 50 mL of 10 X TBE buffer (Gibco) into 950 mL distilled water to total volume of 1 L. Buffer was stored in room temperature.

2.10.6 PCR Purification Buffers (PB, PE, EB Buffers)

Buffers for DNA purification (PB, PE, and EB buffers) were included in QIAquick DNA Purification Kit (Qiagen). Before use, PE wash buffer was mixed with 275 mL 99.97% ethanol (Sigma-Aldrich). Buffers were stored in room temperature.

2.10.7 Ethanol

For 70% ethanol, fourteen milliliters of 99.97% ethanol (Sigma-Aldrich) was added with 6 mL HPLC water (VWR). For 95% ethanol, nineteen mL of 99.97% ethanol was added with 1 mL HPLC water. Both solutions were stored in room temperature.

2.10.8 Sodium acetate, 3M, pH 4.6

Sodium acetate, 3M, pH 4.6 (Applied Biochem) was stored in room temperature

CHAPTER THREE: Pneumococcal Sequetyping Using a Single Primer Pair

3.1 Introduction

3.1.1 Pneumococcal Serotyping

Serotyping the pneumococcus has been an integral part in both pneumococcal research and clinical settings. Serotype-specific serum therapy during the early half of the last century required knowledge of the disease-causing serotype (Finland & Sutliff 1931). More recently, the correct serological identification of *Streptococcus pneumoniae* is crucial in understanding global and regional epidemiological trends, as well as determining vaccine efficacies. Conventional serotyping was based on the Quellung reaction or the Neufeld test, invented by Fred Neufeld in 1902, which described the change in refractive properties of the pneumococcal capsule upon reacting with specific anti-capsular antibodies. In 1931 the Quellung reaction was accepted as a method for differentiating capsular types (Austrian 1976), and to this day it remains to be the "gold standard" for serological classification of the organism because it is relatively easy to perform.

3.1.2 Quellung Reaction

Conventional Quellung reaction serotyping employs pneumococcal antiserum formulated and manufactured by the Statens Serum Institut (Copenhagen) (Henrichsen 1979). Polyclonal rabbit antibodies are served for this diagnostic purpose and have 4 levels of increasing capsular specificity in serum:

1) Omniserum reacts with all serotypes at the time of its development. This can be used as rapid identification method for pneumococci.
2) Pooled sera contains up to 12 pools with each pool containing antibodies specific to groups of serogroups and serotypes.
3) Type or group sera are specific to a particular serotype. For a serogroup with multiple serotypes, group sera will cross-react with serotypes within a group.
4) Factor sera are used to differentiate serotypes within a serogroup.

The main serotyping method was modified by Sørensen (Sørensen 1993), enabling identification of serotypes/groups of 90-95% of pneumococci isolated from blood and

CSF using the chessboard system. The amount of costly sera required for this new method was greatly reduced compared to typing schemes involving a high number of antisera (Lalitha *et al.* 1999). The chessboard system involves 7 existing pooled antisera (A to H) and 5 new pools (P to T) to identify most serotypes of the PPV23. The pooled method was designed so that a serotype/group of the PPV23 serotypes would react with one existing pooled antiserum and one new pooled antiserum, and that the most commonly encountered serotypes would be identified first.

3.1.3 Disadvantages of the Quellung Reaction

The Quellung reaction requires technical expertise as agglutination visualization and interpretation may be subjective. A particular isolate would also need to be tested against a high number of antisera, slowing down identification. Third, a culture step is required, therefore culture-negative clinical isolates may not be detected (Azzari *et al.* 2008). Fourth, there have been reports documenting anti-pneumococcal antibodies reacting with viridans streptococci, leading to false positives (Smith & Washington 1984; Holmberg *et al.* 1985). The main disadvantage remains to be the cost of antisera. A kit of 12 pooled antisera, one mL each vial, costs £1200 in total, and would not be able to differentiate cross-reactive types within a group. To identify types within a group, a full set of factor sera would cost approximately £10,000.

3.1.4 Alternative Serological-Based Serotyping Methods

Most alternative serological-based methods have been developed in the past 3 decades. Their main attempts were to address the aforementioned shortcomings present in conventional serotyping. These methods are summarized in the following sections. While some of these methods are promising, very few of them entered routine use as they introduce other shortcomings.

3.1.4.1 Microscope Precipitin Test

The high cost required in the Quellung reaction was recognised as early as 1938, when Schaub and Reid (Schaub & Reid 1938) developed the microscope precipitin test. This provided improved results when typing pneumococci from cultures, while requiring small amounts of serum. At that time this method also bypassed the need for test tubes and pipettes.

3.1.4.2 Microtiter Plate Agglutination Typing

An agglutination method performed on microtiter plate was developed to ease routine serotyping on large numbers of pneumococci (Kirkman, Fischer & Pagano 1970). Culture was pre-treated with formalin to facilitate apparent agglutination with an unaided eye upon reaction with a factor serum. This method, however, required multiple incubation steps.

3.1.4.3 Capillary Precipitin Typing (CPT)

CPT was designed also to allow large-scale serotyping involving low amounts of immune serum. Type-specific antisera reacted with capsule antigens in glass capillary pipettes (Russell *et al.* 1978). A pipette was dipped into antiserum until capillary action drew up serum to 15 mm in length. The pipette was wiped with tissue paper and dipped into pneumococcal antigen solution of a similar volume. Pipette was then inverted and agglutination is visible for a positive reaction when pipette is placed against a black background with a tungsten light bulb or with a florescent light. There was 100% agreement between the capillary test and the Quellung reaction. However, a positive reaction may take up to 30 minutes to be visible.

3.1.4.4 Counterimmunoelectrophoresis (CIE)

CIE involves electrophoresis on agarose containing paired wells, a cathode one with antigen and the anode with capsular antibodies, and the specific precipitation as they converge along the gel. CIE has been applied to capsular antigen detection in sputum, blood, urine, and CSF samples (Fossieck, Craig & Paterson 1973; El-Refaie & Dulake 1975). Multiple cycles of CIE would be performed, first with omniserum, followed by group and type antisera, to determine the serotype. Holliday *et al.* (Holliday 1981) subsequently devised a one-step method by employing pneumococcal omniserum without the need for group or type antisera. CIE was in wide use for detecting antigens from pneumococci causing invasive diseases during the introduction of the 14-valent pneumococcal vaccine. However, a low proportion of isolates could be serotyped. Further analysis indicated that CIE using conventional barbitol buffer was incapable of serotyping capsules with neutral charge such as 7F, 7A, 14, 33F, 33A and 37, and these types could be correctly identified using *m*-carboxylphenylboronic acid buffer instead (Henrichsen, Berntsson & Kaijser 1980).

3.1.4.5 Co-Agglutination (COA)

In this method, serotype-specific antibodies are bound and stabilised on their F'c fragments to protein-A of *S. aureus* (Kronvall 1973). The antibody-protein-A complex is then mixed with the pneumococcal strain on a glass slide (Smart 1986). The multiple serotype-specific F'ab regions present allow for the strong amplification of the agglutination signal. A positive signal could be visible within seconds. COA has been reported to be superior to CIE, being able to identify serotypes with neutral charges (Trollfors *et al.* 1983), is as accurate as Quellung (Smart 1986), and is superior to CIE and Quellung in simplicity and speed in serotyping. The chessboard system (Section 2.6.4) was also modified for COA typing (Lalitha *et al.* 1999).

3.1.4.6 Latex Agglutination (LA)

LA was first employed to detect β-streptococci, and subsequently for capsular typing of pneumococci (Browne, Miegel & Stottmeier 1984). In this method, antisera are coated with latex particles and are mixed with capsule antigen of an isolate. Visualization of agglutination with an unaided eye is indicative of a positive reaction, while the presence of a milky suspension suggests a negative reaction (Lafong & Crothers 1988). Results from LA was reported to be in agreement to Quellung serotyping and coagglutination, and may be a cost-effective alternative for serotyping in developing countries (Kaldor, Asznowicz & Dwyer 1988; Lafong & Crothers 1988; Lalitha *et al.* 1996). LA may also be applied directly to body fluids (Kaldor, Asznowicz & Dwyer 1988; Sridharan *et al.* 1994). Statens Serum Institut (Copenhagen) manufactures latex-coated serotyping kit (Pneumotest-Latex), based on the chessboard method (Slotved *et al.* 2004). This method has been shown to produce false-positive results with viridans streptococci (Browne, Miegel & Stottmeier 1984; Slotved *et al.* 2004).

3.1.4.7 Fluorescence Antibody Technic (FAT)

FAT involves the detection of capsular antigen by rabbit immunoglobulin, and FITC-conjugated, anti-rabbit goat immunoglobulin. The antigen and antibodies are mixed onto films and observed under the fluorescent microscope. While the method may be more time-consuming than the Quellung reaction, the main advantage of FAT is that fluorescence eases visualization, and it is more sensitive in identifying serotypes in various body fluids with low bacterial titre or with small capsules (Wicher *et al.*

1982). However, cross-reactivity with other α- and β-haemolytic streptococci and *Klebsiella pneumoniae* was reported.

3.1.4.8 Dot Blot ELISA Assay (DBEA)

The DBEA was developed by Fenoll *et al.* (Fenoll *et al.* 1997) and was being used routinely by the group at the time of their publication in addition to the Quellung test. A heavy pneumococcal suspension in saline was spotted onto nitrocellulose membranes and dried at room temperature, followed by membrane blocking with 5% skim milk in PBS. After washing type or group antiserum was added to the membrane and incubated on a shaking platform. Goat anti-rabbit immunoglobulin G conjugated with peroxidase is added. When mixed with substrate 4-chloronaphthol, a positive reaction is seen as a dark colour surrounded by a halo. This method is reported by the group to be specific, rapid, and economical, but identification of types within a serogroup may be inconclusive.

3.1.4.9 Immunoblot ELISA Assay (IBEA)

The IBEA is a variation of DBA in which nitrocellulose membrane is placed directly on top of blood agar plate (Bronsdon *et al.* 2004). This allows for detection of multiple serotypes from nasopharyngeal samples, where other serotypes may be low in prevalence. Bronsdon *et al.* (Bronsdon *et al.* 2004) acknowledges the tedious nature of blot analysis and culture, which presents obstacles for rapid diagnostic use and skews proportions of co-colonizing serotypes in multiple colonizations.

3.1.4.10 Flow Cytometry Assay

The assay allows simultaneous detection of up to 15 serotypes, and also enabled to differentiation of cross-reactive serotypes (Park, Briles & Nahm 2000). Latex beads of variable sizes with red-fluorescent dye were coated with pneumococcal capsule antigen. This was then mixed with bacterial lysate in a microplate with rabbit pneumococcal antiserum pool P-T from Statens Serum Institut (Copenhagen), along with a fluorescein-conjugated anti-rabbit antibody. A correct match is the serotype in the lysate that effectively reduces the green fluorescence given by the antibody and the antigen-coated latex beads. Despite the small sample size, this method showed 100% agreement to conventional typing, including cross-reactive types. The serotype coverage of this method was later expanded to cover 36 serotypes with incorporation

of a multiplex PCR reaction (Yu *et al.* 2005; Yu *et al.* 2008), and had been validated on clinical samples from different continents (Lin *et al.* 2006). With the use of Luminex® technology, flow cytometry-based serotyping has recently been shown to be efficient and rapid in determining most of the known pneumococcal serotypes and detection of multiple serotypes (Yu *et al.* 2011).

3.1.5 DNA-based Serotyping

Serological methods mentioned above require antigens and antibodies in sufficient amounts to enable detection. With the recent advances in molecular biology and bioinformatics, DNA-based serotyping methods have been pursued. Molecular methods drastically increase sensitivity of serotyping, allowing identification of culture-negative pneumococci found in clinical samples (Saha *et al.* 2008; Tarragó *et al.* 2008). A recent comparison between DNA-based and Quellung serotyping revealed that colonisation rates in healthy children may be two-fold greater if molecular serotyping method was employed in place of conventional serotyping (Ogami *et al.* 2010). This is also important in the analysis of multiple colonization events (Chapter 4), as multiple serotypes can only be detected in culture by sampling multiple colonies on agar (Huebner *et al.* 2000a; Bronsdon *et al.* 2004) or by enrichment liquid broth (Kaltoft *et al.* 2008). For serotype identification, DNA-based typing methods target the serotype-specific genes of the capsular polysaccharide synthesis (*cps*) operon (Kong & Gilbert 2003; Kong *et al.* 2005; Pai, Gertz & Beall 2006; Turner *et al.* 2011). DNA-based methods may be applied directly on clinical samples, bypassing a culture step. Compared to culture serotyping, PCR-based methods are much less laborious and efficient.

3.1.5.1 Multiplex/Multistep PCR-Based Serotyping Methods

A plethora of DNA-based serotyping methods targeting serotype-specific genes, as the presence of these regions clearly identifies the serotype. These methods are briefly summarized below.

3.1.5.2 Multiplex PCR in Determining Nine Common Serotypes

This method aimed to target serotypes covered in the proposed 11-valent conjugate vaccine (Brito, Ramirez & De Lencastre 2003). Multiple primers specific to serotypes 1, 3, 4, 6B, 14, 18C, 19F, 19A, and 23F were put into six different multiplex PCR

reactions. Differentiation of serotypes within serogroup 6 and 18 were shown to be problematic, however the use of expensive antibodies can be reduced after serogroup determination of these strains.

3.1.5.3 Semiautomated Multiplex PCR

The study designed by Lawrence *et al*. (Lawrence *et al.* 2003) employs eight primer pairs to detect serotypes 1, 3, 14, 19F, 23F, and serogroups 6, 19, and 23. Of 93 pneumococci tested in the study, all within the targeted serotypes and serogroups were identified correctly and reproducibly. This method was also species-specific, with no amplification products from related streptococci. However, not all primer pairs are put into the same reaction, and with an unknown pneumococcus three separate multiplex PCR reactions may be required for serotype identification.

3.1.5.4 Combination of *cpsA*-*cpsB* polymorphism and serotype-specific *wzy* and *wzx*

Kong and Gilbert (Kong & Gilbert 2003) designed a method that is able to differentiate up to 51 serotypes based on sequencing capsular regulatory genes *cpsA* and *cpsB*, as well as serotype-specific PCRs targeting the polymerase (*wzy*) and flippase (*wzx*) genes. The sequencing aspect of this work will be discussed below (Section 3.1.5.14), however PCR branch of this method is able to differentiate 12 different serotypes using more than 20 primer pairs (some serotypes require more than one primer pair per gene). PCR alone is not able to differentiate between members of serogroup 6, 18 (which can be resolved by the sequencing branch of this method), and is problematic differentiating between members of serogroup 33 except for 33B. This method eventually expanded to cover all 90 serotypes known at that time (Kong *et al.* 2005) and was tested on a large number of pneumococci (a total of 708 pneumococci from the two studies). An additional 26 new primer pairs targeting the serotype-specific genes enabled differentiation of serotypes in which sequencing results were ambiguous. However even with the expanded repertoire of primer pairs, members of serotypes 6A and 6B proved difficult to differentiate by this method.

3.1.5.5 Multiplex Nested PCR Detecting Serotypes 3, 14, 19F, 23F

This method uses dot-blot probe to detect amplification by employing labelled probe-linked outer and inner primers to enhance detection of pneumococci (Rubin & Rizvi 2004). This study tested more than 50 serotypes and 28 non-pneumococcal strains,

and has proven the method to be specific both at the species and the serotype level (for the serotypes targeted). Of the pneumococci correctly serotyped, the sensitivity of this method is extremely high (2-10 fg or one bacterial cell from exponential phase), which is also shown by the detection of pneumococci in culture-negative samples. The major drawback of this method is that only four serotypes are covered.

3.1.5.6 One-Step Multiplex PCR to Detect PCV7 Serotypes

Serotypes covered in PCV7 can be correctly identified using this method bypassing the need multiple PCR reactions (O'Halloran D & Cafferkey 2005). Instead, all seven primer pairs are included in this procedure to identify PCV7 types. This method had been shown to be applicable to clinical samples from invasive diseases (Azzari *et al.* 2008). However, similar to other studies, differentiation between serotypes PCV-related type 6A and 6B was proved to be difficult.

3.1.5.7 Differentiation of Serotypes 6A and 6B by Pyrosequencing

Pai *et al.* (Pai, Limor & Beall 2005) have identified that despite the high homology shared between serotypes 6A and 6B, the two types can be differentiated by pyrosequencing based on a SNP at codon 195 of the *wciP* gene, encoding a rhamnosyl transferase (Mavroidi *et al.* 2004). There was 100% congruence between SNP and serotype in that region on over 200 pneumococci tested. This method is likely to be used alongside with other serotyping method as this method solely differentiates between these two serotypes.

3.1.5.8 Multistep Multiplex PCR

Pai's group (Pai, Gertz & Beall 2006) also devised a multiplex, multistep PCR reaction that can differentiate 17 serotypes, and can be employed directly on clinical invasive and carriage samples with higher sensitivity compared to culture methods (Azzari *et al.* 2008; Saha *et al.* 2008; Antonio *et al.* 2009; Moore *et al.* 2010). More than 20 primer pairs are included in seven sequential multiplex PCR reactions. The order of the reactions was optimized originally by Pai *et al.* (Pai, Gertz & Beall 2006) to allow early detection of serotypes commonly found in the United States (Fig. 3.1). However, the ordering of the reactions can be altered to account for geographical differences in serotype prevalence (Dias *et al.* 2007; Morais *et al.* 2007; Saha *et al.* 2008; Iraurgui *et al.* 2010; Njanpop Lafourcade *et al.* 2010). Different multiplex

reactions could also be combined to decrease workload. With this method, up to three serotypes were detected to be simultaneously present in nasopharyngeal secretions (Antonio *et al.* 2009). However, rare serotypes would require repetitive PCR reactions before diagnosis, and more than half of known serotypes would still require conventional typing, and serotypes 6A and 6B could not be differentiated by this method alone, and require the aforementioned pyrosequencing method (Pai, Limor & Beall 2005) or conventional serological typing.

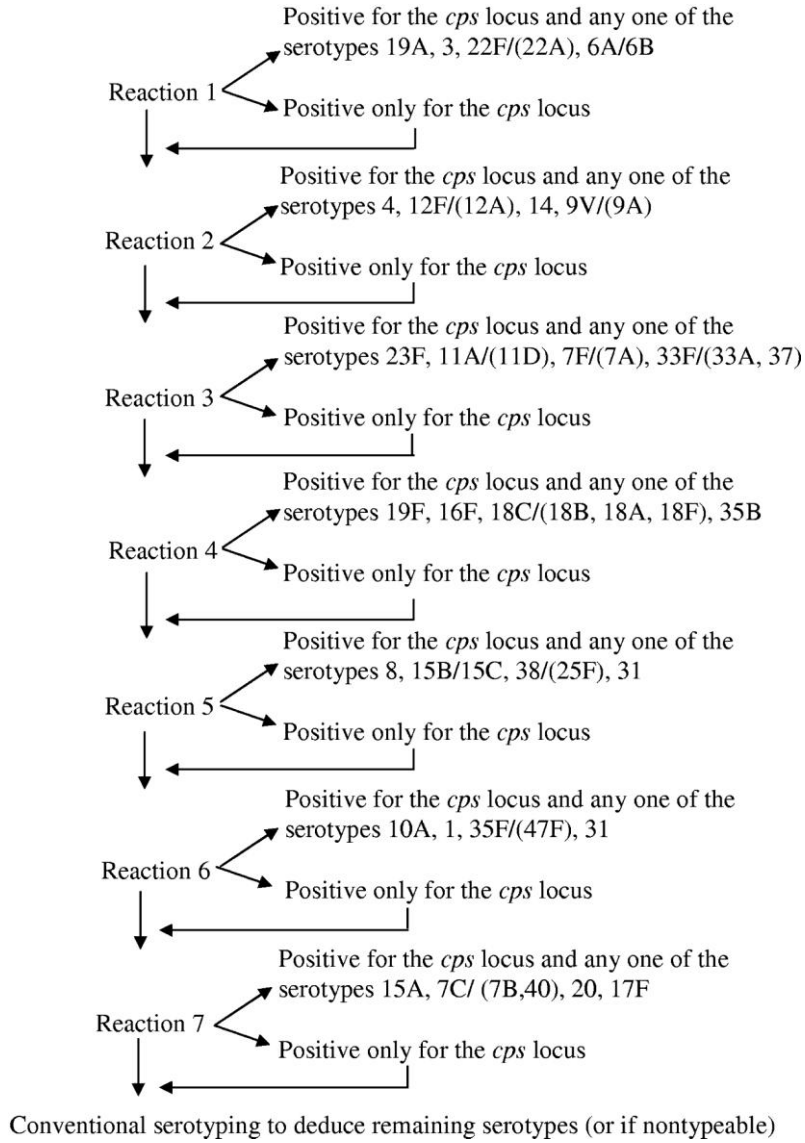


Fig. 3.1. Methodology of Multistep/multiplex PCR reaction. Multiplex PCR reactions performed sequentially to determine one of 17 serotypes. Method designed by Pai *et. al.* (Pai, Gertz & Beall 2006). Samples with no PCR amplification following seven reactions require conventional serotyping. Reaction orders and serotype combinations can be modified based on local epidemiology. Figure from Pai *et al.* (Pai, Gertz & Beall 2006). Journal of Clinical Microbiology authorizes the reuse of journal material including figure shown for academic purposes

3.1.5.9 Multiplex PCR for Pneumococcal Isolated in Otitis Media Patients

The method designed by Billal *et al.* (Billal *et al.* 2006) allows detection of serotypes 1, 3, 4, 14, 19F, 23F, and the emerging serotype 19A, as well as serogroups 6, 18, 19, and 23. However, band sizes produced by this method are similar across serotypes (all amplicons within 100-480 bp long, making differentiation of serotypes by this method difficult.

3.1.5.10 Reverse-Line Blot Hybridization Serotyping

The protocol for this method is described thoroughly by Kong *et al* (Kong & Gilbert 2006). Previously used to detect serotypes of GBS (Kong, Ma & Gilbert 2005), this procedure was then applied to identification of pneumococcal serotypes, targeting serotype-specific gene *wzy* and regulatory gene *wze* (*cpsD*). In brief, oligonucleotide probes labelled with amine is anchored covalently to a nylon membrane. PCR products containing biotin is then hybridized to probes and detected as light emission by introducing peroxidase-containing streptavidin and electrochemiluminescence substrates (Fig. 3.3). This method was tested to detect serotypes covered in the PPV23 (Kong *et al.* 2006), and subsequently expanded its coverage to include all serotypes (Zhou *et al.* 2007). However, cross-reactions were seen for up to 30 different serotypes, most of which were seen between types of a serogroup. Nonetheless, this method would be able to perform serotyping of over 90% of strains encountered, reducing the need and subsequent cost for specific antisera.

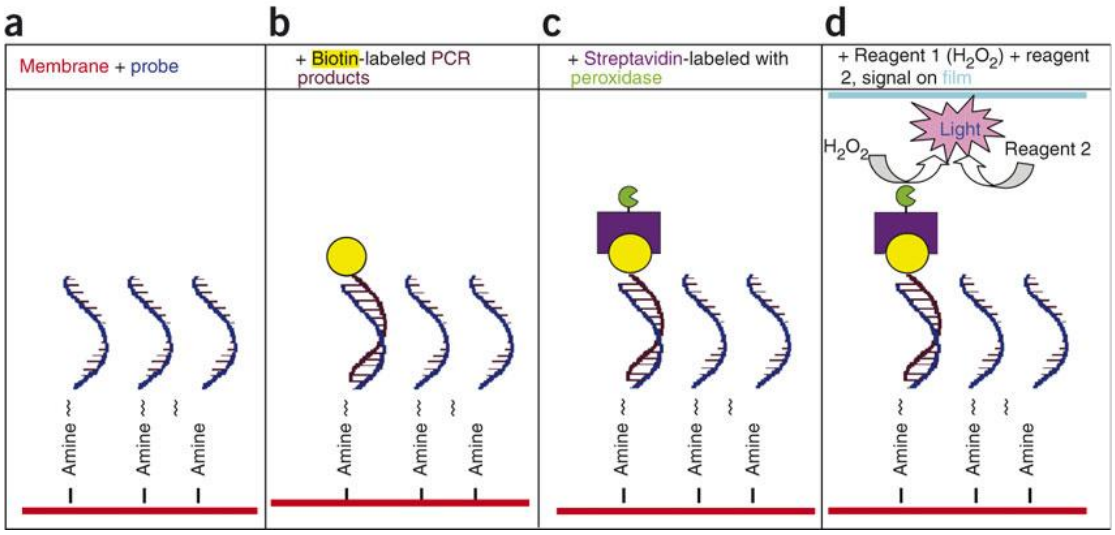


Fig 3.2. Procedures of reverse-line blot hybridization method targeting pneumococcal serotypes based on *wzy* and *wze*. A) Amine-containing serotype-specific oligonucleotide probes covalently bound to nylon membrane. B) Multiplex PCR products containing biotin hybridizes to probes. C) Peroxidase-containing streptavidin binds to the biotin on PCR products. D) Addition of reagents leads to emission of light upon reaction with, which is indication of presence of PCR products. Figure from Kong *et. al.*(Kong & Gilbert 2006) with permission granted from Nature Publishing Group via RightsLink Copyright Clearance Center.

3.1.5.11 Multiplex Real-Time PCR

By targeting different serotype-specific genes, twelve serotypes (1, 3, 4, 5, 6A, 6B, 8, 14, 15B/C, 19A, 23A, and 23F) can be differentiated by this real-time PCR approach with high sensitivity (detection level of 12-24 fg or five to ten bacterial cells per reaction) (Tarragó *et al.* 2008). Culture-negative samples could be detected to contain pneumococcal DNA from pleural fluids of children suffering from parapneumonic empyema. This method was promising as a tool for detecting pneumococci in cases of parapneumonic empyema, which conventional cultural detection methods have proven insensitive, but numerous related serotypes within a group (such as serogroups 7, 9, 18, and 19) could not be differentiated.

3.1.5.12 Restriction Fragment Length Polymorphism (RFLP)

RFLP serotyping methods involve the *cpsAB* regulatory region (Lawrence, 2000) or the entire *cps* locus (Batt, 2005). Targeting c*psA* and *cpsB* showed modest association between restriction pattern and serotype, with exceptions to a number of serotype 6A strains with identical banding patterns to some serotype 6B pneumococci. This

method employed multiple restriction enzymes, thus potentially complicating analysis. Therefore, this method is superior to conventional typing only from an economic point of view. Different sets of restriction patterns were also seen for isolates within a serotype (Lawrence, 2000). By including the entire *cps* region for restriction analysis, differences in serotype-specific genes can also be reflected as different banding pattern (Batt, 2005). RFLP of entire capsulation locus from *dexB* to *aliA* using a single endonuclease (*Hin*FI) has revealed 46 restriction patterns that correlated well with serotypes with reproducibility. With the exception of serotype 6B, multiple strains of the same serotype were observed to share over 90% similarity in banding patterns when displayed as a dendogram (Batt, 2005). This method may be a promising alternative, however amplification of a fragment of up to 30kb is technically demanding and time-consuming.

3.1.5.13 Microarray

Wang *et al.* (Wang *et al.* 2007) designed a microarray method using probes targeting serotype/group-specific regions in *wzy* (*cpsD* in serotype 3) to identify 43 serotypes including the 23 included in PPV23. This method was shown to be species-specific however closely-related species such as *S. mitis* and *S. oralis* were not included in the analysis. In another recent study, serotype/group-specific glycosyltransferase genes were used as targets for serotyping a small number of pneumococci of PPV23 serotypes (Tomita *et al.* 2011a). Unfortunately, vaccine-related serotypes may also be identified as a vaccine type, as in the case of a 23A pneumococci being typed as 23F by this method. With the availability of Wellcome *cps* sequences, microarray could be used to differentiate the majority of the serotypes, with the ability to detect serotypes with low abundance in co-colonization samples (Turner *et al.* 2011).

Recently, serotype coverage by microarray has been expanded to identify most of the 93 serotypes known to date. The sensitivity of microarray was also demonstrated to be superior to that of culture method, with a higher rate of co-colonization and extent of serotype diversity detected within a co-colonization sample (Chapter 4). However, the cost of setting up a microarray limits its widespread use, especially in developing countries where pneumococcal disease burdens are high.

3.1.5.14 Sequencing-Based Serotyping Methods Targeting *cpsAB*

The targeting of serotype-specific genes in multiple PCR means that a large number of primers would be required for appropriate coverage of serotypes. To enhance efficiency and to simplify set-up process, the design of a universal primer would be preferable. A sequence-based serotyping method took advantage of the polymorphisms within *cpsA* and *cpsB* (Jiang, Wang & Reeves 2001; Kong & Gilbert 2003; Kong *et al.* 2005). Based on *cps* sequences of 11 serotypes (1, 2, 4, 6B, 8, 14, 18C, 19F, 19A, 23F, 33F), a variable region flanked by conserved sequences was detected from 950 bp downstream of *cpsA* to 302 downstream of *cpsB*. A combination of *cpsAB* sequencing and *wzx*/*wzy*-specific PCR was employed for pneumococcal serotyping (Kong & Gilbert 2003; Kong *et al.* 2005). Sequence analysis of over 700 *cpsAB* sequences revealed 138 sequence types between 90 serotypes, with most of the sequence types associated with a single serotype. The use of *wzx* and *wzy*-specific PCR assisted in serotyping ambiguous sequence types shared by multiple serotypes (e.g. 23F/A) (Kong & Gilbert 2003). Combining *cpsAB* sequencing and serotype-specific PCR, this method could identify 95% of pneumococci at least to the serogroup level (Kong *et al.* 2005).

3.1.5.15 Weaknesses of *cpsAB* Partial Sequencing

The sequence-based typing method described above suffers from some drawbacks. Similar to previous reports, serotypes 6A and 6B remain to be difficult to differentiate. In addition, with the new identification of serotypes 6C and 6D, the resolving power of this method to these new serotypes is unknown. Other serotypes (10A, 13, 14, 15B, 17F, 18B, 20, 22F, 23F, 23A, 25F, 33F, 33A, 35B, 38, 42) provided ambiguous results, and only five of these serotypes (6A, 6B, 18B, 33F, 33A) were serogroup-specific. Most importantly, the primer-binding region was located based on *cps* sequences of only 11 serotypes available at the time (Jiang, Wang & Reeves 2001). With the recent sequence analysis of *cps* operon of over 90 serotypes, more heterologous primer-binding regions may be identified by aligning *cps* sequences from these serotypes.

3.1.6 Aim

At present, the availability of *cps* operon sequences of 90 serotypes from Wellcome Trust (Bentley *et al.* 2006), in addition to that of serotypes 6C (Park *et al.* 2007) and 6D (Bratcher *et al.* 2010) allows us to revisit the potential of a sequence-based sequetyping system. Specifically, we aim to re-identify a heterologous region across all known serotypes, flanked by conserved sequences for binding of a single primer. Therefore, our aim is to locate this region using *in silico* multiple alignment method. Upon identification of this region, a common primer pair was tested on reference and clinical pneumococcal strains.

3.2 Materials and Methods:

3.2.1 Sequence Alignment of *cps* of PPV23 Serotypes

An *in silico* interrogation was performed by Dr. Kevin Bryson (UCL Bioinformatics) on the *cps* locus sequences made available by the Wellcome Trust (Bentley *et al.* 2006). Alignment of *cps* of pneumococcal serotypes constituting the PPV23 was performed to identify 500 to 1100bp variable regions that are bordered by conserved primer sites for PCR amplification that could differentiate between serotypes. Alignment of *cps* sequences for the 23 vaccine serotypes was performed using MAFFT version 6.624b online interface (http://mafft.cbrc.jp/alignment/server/) using the iterative refinement method (NS-i). ClustalX 2.0.10 was used to analyze the alignment generated by MAFFT and generate a q-scores file, which contains information regarding the conservation of each nucleotide position along the multiple alignment; each nucleotide position was scored (i.e. consensus value for each nucleotide position) according to its conservation across all the 23 serotype sequences. A high conservation score would thus suggest a suitable common primer binding site due to its high sequence conservation. The q-scores file was then entered into an algorithm (PrimerFinder algorithm) to locate conserved primer-binding sites. After identification of an optimal primer pair based on maximum conserved binding sites and sequence differentiation between all 23 serotypes, the region to be amplified by this primer pair was analyzed *in silico* for 92 published serotypes (see below).

3.2.2 "Primer Finder" Algorithm for Selecting Optimal Primer-Binding Sites

The algorithm was developed by Dr. Kevin Bryson to identify potential conserved primer-binding sites, as measured from multiple alignment conservation q-score measurements generated from ClustalX. The optimal binding sites would be conserved across all 23 serotypes, flanking sequences that are unique in all serotypes. The PrimerFinder algorithm allows user to set the desired primer length, the amplicon length, and the number of nucleotide mismatches allowed on each primer-binding site. Thus, the initial primer length was set to 20 nucleotides, the amplicon length was set from 500 bp to 1,100 bp, and the number of nucleotide mismatches was set to two nucleotides per binding site. The algorithm examined each alignment position in turn within the multiple sequence alignment of the *cps* operon for all 23 sequences, searching for putative forward and reverse primers that obey the set values and whose amplicon sequence can discriminate the largest number of serotypes.

Each alignment position in turn was considered as a putative start site for a forward primer (Fig. 3.4, Step 1). The length of the primer was then extended one base pair at a time to obtain the longest forward primer while retaining the maximum number of matching sequences (Fig. 3.4 Step 2). This was followed by a search for the most optimal reverse primer site by searching the downstream alignment positions at a distance from 500 to 1100 bp from the upstream primer site. The primer extension approach for each putative reverse primer was repeated (Fig. 3.4 Step 3). For each putative primer pair examined, the internal amplified sequences across all the matching sequences were analyzed (Fig. 3.4 Step 4). Serotypes that could be putatively differentiated by unique sequences between the primer pair were then recorded. The serotypes with identical sequences in this region were recorded as ambiguously identified. The interrogation was repeated until all possible primer pairs along the alignment were considered. The results were recorded and ranked by the total number of serotypes (in decreasing order) each primer pair could identify.

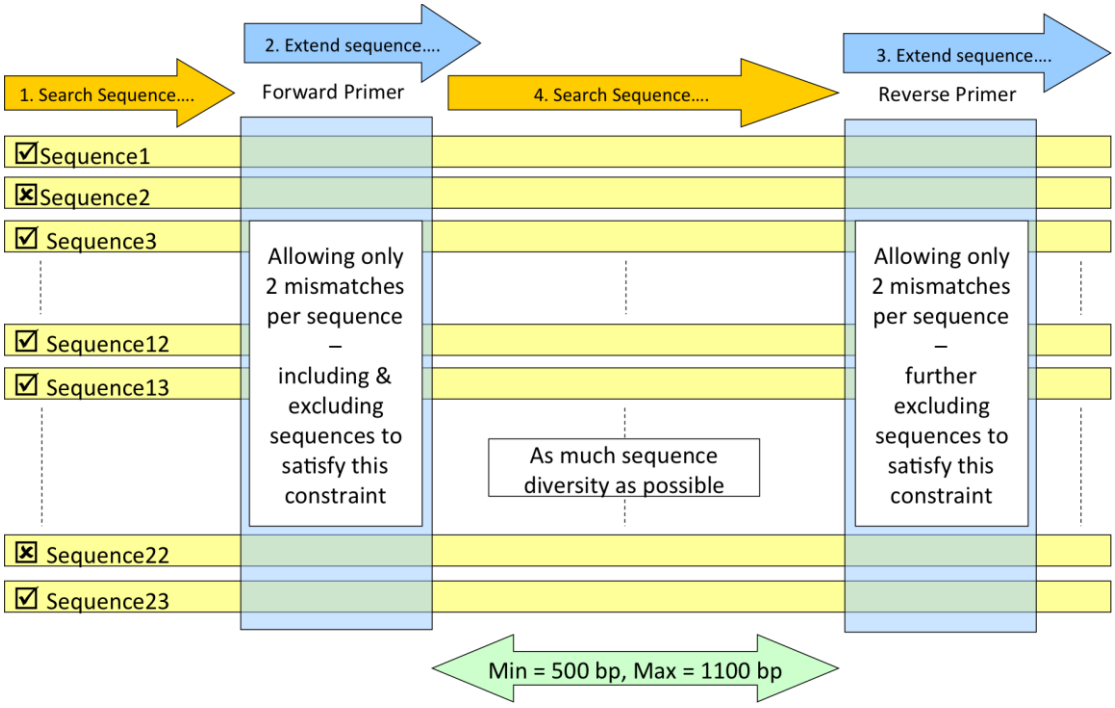


Fig. 3.3. Schematic representation of searching approach used by Primer Finder algorithm. The forward primer is identified within the multiple sequence alignment (allowing 2 mismatches per primer and including as many sequences as possible), then the search finds a reverse primer between 500 and 1100 bases downstream of the first, again satisfying the constraints and also having as much sequence diversity within the internal amplicon region. Steps 1 to 4 are described in Materials and Methods.

The algorithm identified four potential primer-binding sites (Table 3.1) differing in amplicon lengths (between ~600 to ~1000 bp) and number of nucleotide variation (1-3 nucleotides) in amplified region between the 23 *in* silico-tested serotypes. As primer pair 1 provides the longest amplicon and the highest number of nucleotide variations between serotypes, this primer pair chosen by Mr. Marcus Leung to be tested on bacterial strains for this study (see below). Primers were designed in Primer3 (http://frodo.wi.mit.edu/primer3/) by using template sequence of strain 519/43 (serotype 1, accession number CR926497) and the 5' and 3' ends of the heterologous region identified above as respective primers. Single nucleotides were substituted on the non-extending ends of the primers to adjust for similar annealing temperature and minimize self-annealing. Sequences for primers used (*cps-fw* and *cps-rv*) are indicated in Table 2.6.

3.2.3 Bacterial Strains and PCR

Please refer to Appendix Table A1 for strains included in this study. Conventional serotyping using pooled and typed antisera was performed by Mr. Marcus Leung as described in Section 2.6.4. Please refer to Table 2.5 for standard PCR master mix. For testing species-specificity of sequetyping, an internal control PCR targeting 16S ribosomal RNA gene was employed (Woo *et al.* 2001). The primers and Tm for the control PCR are indicated in Table 2.6. Gel electrophoresis, DNA purification, cycle sequencing, ethanol precipitation, and sequence analysis were performed as outlined in Sections 2.8.3 to 2.8.8.

3.2.4 Interrogation of Query Sequence to GenBank Database

Raw sequence data were trimmed by Mr. Marcus Leung to include intact *cpsB*. For a given query sequence, the serotype associated with a result sequence bearing the highest BLAST bit score was the identified serotype by our novel method. Should the sequence with the top BLAST score belong to the same serogroup, the result is termed "serogroup-specific." In cases where multiple serotypes of different serogroups share the highest BLAST score, results are treated as "ambiguous." A misidentified result is one where the top BLAST score corresponds to a sequence of a different serogroup.

Table 3.1. Putative primer pairs identified from alignment of *cps* from vaccine serotypes

Pair[a]	Amplicon Length (bp)[b]	Primer[c]	Location[d]	Sequence (5'-3' on coding strand)	Minimum number of nucleotide differences between unique amplicon sequences[a]
1	1058	Forward	4414	ATGCCAGACAGTAACCTCTAT	3
		Reverse	5471	AATCAGGACTTGCAGGCAGG	
2	1056	Forward	3772	CCTATTAGTTCGGTGTCGCGATCAGATGTCAATATC	2
		Reverse	4827	AATGATAGTCGTTATGCCTTGAT	
3	832	Forward	4414	ATGCCAGACAGTAACCTCTAT	2
		Reverse	5245	ATCAACTAATTTAGGAGAAA	
4	667	Forward	4805	AATGATAGTCGTTATGCCTTGAT	1
		Reverse	5471	AATCAGGACTTGCAGGCAGG	

[a] Pair 1 used for *in vitro* sequetyping due to identification in all 23 serotypes with highest number of nucleotide differences (i.e. all 23 serotypes differ by at least three nucleotides)

[b] *In silico* anticipated primer pair one of a length of 1058 bp, however primer used in *in vitro* sequetyping has added three nucleotides to 5' end of forward primer, therefore the actual amplicon length is 1,061 bp

[c] Pair 2 reverse and Pair 4 forward are identical in location of the coding strand

[d] Numbers based on nucleotide position on accession number CR926497.

3.3 Results

3.3.1 *In Silico* Identification of Optimal Primer Pair

In silico evaluation has identified a region with sequence variation between the 23 vaccine serotypes with a 1,061 bp amplicon that is 1,351 bp downstream of *cpsA* to 224 bp downstream of *cpsC* (nucleotide position 305,125 to 306,185 on the G54 genome, GenBank accession number NC011072). Subsequent *in silico* analysis of the 92 serotype sequences revealed that with each primer binding site set at a mismatch of no more than two nucleotides, up to 84 of the 92 serotypes could yield an amplicon (Table 3.2).

As some of the amplifiable serotypes share identical interceding sequences (e.g. sequences for serotypes 2 and 41A, as well as 7B and 40, Table 3.2), a subset of 54 of the 84 amplifiable serotypes would be differentiated by *in silico* sequence analysis of the amplicon (Table 3.2). However, due to the sequencing chemistry, the entire 1,061-bp region between the primers is unlikely to be sequenced successfully; approximately 200 bp at each end may yield suboptimal sequence data. *In silico* analysis was therefore repeated to determine the serotypes that could be sequetyped when the central 732 bp region of the *cpsB* amplicon (Table 3.2). This revealed an additional 8 serotypes (6C, 6D, 7F, 7A, 17F, 18B, 18C, 33C) that would not be identified by *cpsB* alone. Of these eight serotypes, three pairs could be sequetyped to the correct serogroup, namely, 6C/6D, 7F/7A, and 18B/18C. Serotypes 17F and 33C have identical *cpsB* sequences and were not predicted to be differentiated from each other.

Serotypes 25A, 25F and 38 were predicted to be non-amplifiable as the forward primer binding site has six mismatches. However, subsequent *in silico* sequence analysis of these three serotypes revealed multiple weaker reverse binding sites with at least 10 nucleotide mismatches that are present approximately 2,000 bp downstream of the forward primer, situated within the glycosyl transferase, *wcyA* gene.

Table 3.2. *In silico* analysis of pneumococcal serotypes that can be amplified and differentiated by sequetyping based on sequence diversity of *cpsB*.

Serotype (n = 92)[b]	Amplification possible [a,c]	Differentiable by *in silico* analysis [a, c, d]	Differentiable by *cpsB* PCR and sequencing? [a, c, d]	Included in *in vitro* testing [a]
1	Y	Y	Y	Y
2	Y	(2/41A)	(2/41A)	N
3	Y	Y	Y	Y
4	Y	Y	Y	Y
5	Y	Y	Y	Y
6A	Y	Y	Y	Y
6B	Y	Y	Y	Y
6C	Y	Y	(6C/6D)	Y
6D	Y	Y	(6C/6D)	N
7F	Y	Y	(7A/7F)	Y
7A	Y	Y	(7A/7F)	N
7B	Y	(7B/40)	(7B/40)	N
7C	Y	Y	Y	Y
8	Y	Y	Y	Y
9A	Y	(9A/9V)	(9A/9V)	N
9L	Y	Y	Y	N
9N	Y	Y	Y	Y
9V	Y	(9V/9A)	(9V/9A)	Y
10F	Y	(10F/10C)	(10F/10C)	Y
10A	Y	Y	Y	Y
10B	Y	Y	Y	N
10C	Y	(10C/10F)	(10C/10F)	N
11F	N	Unknown	Unknown	N
11A	Y	(11A/11D/18F)	(11A/11D/18F)	Y
11B	Y	(11B/11C)	(11B/11C)	N
11C	Y	(11B/11C)	(11B/11C)	N
11D	Y	(11A/11D/18F)	(11A/11D/18F)	N
12F	Y	(12F/44)	(12F/44)	Y
12A	Y	Y	Y	N
12B	Y	Y	Y	Y
13	Y	(13/20)	(13/20)	N
14	Y	Y	Y	Y
15F	Y	Y	Y	N
15A	Y	Y	Y	Y
15B	Y	Y	Y	Y
15C	Y	Y	Y	N
16F	Y	Y	Y	Y
16A	Y	Y	Y	N
17F	Y	Y	(17F/33C)	Y
17A	Y	(17A/34)	(17A/34)	Y
18F	Y	(11A/11D/18F)	(11A/11D/18F)	N
18A	Y	Y	Y	N
18B	Y	Y	(18B/18C)	Y
18C	Y	Y	(18B/18C)	Y
19F	Y	Y	Y	Y
19A	Y	Y	Y	Y
19B	Y	Y	Y	N
19C	Y	Y	Y	N
20	Y	(20/13)	(20/13)	Y
21	Y	Y	Y	Y
22F	Y	(22F/22A)	(22F/22A)	Y
22A	Y	(22A/22F)	(22A/22F)	Y
23F	Y	Y	Y	Y
23A	Y	Y	Y	N

23B	Y	Y	Y	Y
24F	Y	Y	Y	Y
24A	Y	Y	Y	N
24B	Y	Y	Y	Y
25F	N	Unknown	Unknown	N
25A	N	Unknown	Unknown	N
27	N	Unknown	Unknown	Y
28F	Y	Y	Y	Y
28A	Y	Y	Y	Y
29	N	Unknown	Unknown	Y
31	Y	Y	Y	Y
32F	Y	(32F/32A)	(32F/32A)	N
32A	Y	(32A/32F)	(32A/32F)	N
33F	Y	(33F/33A/35A)	(33F/33A/35A)	Y
33A	Y	(33A/33F/35A)	(33A/33F/35A)	N
33B	Y	Y	Y	Y
33C	Y	Y	(17F/33C)	Y
33D	Y	Y	Y	Y
34	Y	(34/17A)	(34/17A)	Y
35F	Y	(35F/47F)	(35F/47F)	N
35A	Y	(35A/33F/33A)	(35A/33F/33A)	Y
35B	Y	(35B/35C)	(35B/35C)	Y
35C	Y	(35C/35B)	(35C/35B)	N
36	Y	Y	Y	Y
37	Y	Y	Y	N
38	N	Unknown	Unknown	N
39	N	Unknown	Unknown	N
40	Y	(7B/40)	(7B/40)	N
41F	Y	Y	Y	N
41A	Y	(2/41A)	(2/41A)	N
42	Y	Y	Y	N
43	N	Unknown	Unknown	N
44	Y	(12F/44)	(12F/44)	N
45	Y	Y	Y	N
46	Y	Y	Y	N
47F	Y	(35F/47)	(35F/47)	N
47A	Y	Y	Y	N
48	Y	Y	Y	N
Total	**Y = 84/92** **N = 8/92**	**Y=54/92** **Ambiguous = 30/92** **Unknown=8/92**	**Y = 46/92** **Ambiguous = 38/92** **Unknown = 8/92**	**Y = 48/92** **N = 44**

[a] Y = yes, N = no.

[b] Serotypes of PPV23 and PCV13 are indicated in bold.

[c] Serotypes that are amplifiable by PCR but give ambiguous results are indicated as two or more serotypes having the same top BLAST bit scores.

[d] Serotypes deemed non-amplifiable in initial *in silico* analysis are not included in the subsequent tests on identifiable serotypes, therefore resolution power is unknown for these serotypes.

3.3.2 Species Specificity of Sequetyping

Other bacterial species, including a range of streptococci and common mucosal colonizers were selected to test the specificity of the selected pair of primers. An amplicon of the expected size of 1,061 bp was detected only in *S. pneumoniae* ATCC 49619 (serotype 19F). Amplification of rRNA gene as a positive control in non-pneumococcal strains confirmed the presence of a DNA template in these samples (Table 3.3).

Table 3.3. The specificity of pneumococcal sequetyping against other bacterial species

Organism	Type Strain[a]	16S rRNA gene[b]	Expected 1,061 bp band
Streptococcus pneumoniae	ATCC 49619	+	+
Streptococcus gordonii	NCTC 7865	+	-
Streptococcus anginosus	NCTC 10713	+	-
Streptococcus intermedius	NCTC 2227	+	-
Streptococcus mitis	NCTC 12261	+	-
Streptococcus mutans	NCTC 10449	+	-
Streptococcus parasanguinis	NCTC 55898	+	-
Streptococcus salivarius	NCTC 8618	+	-
Streptococcus sanguinis	NCTC 7863	+	-
Streptococcus sobrinus	NCTC 12279	+	-
Streptococcus cristatus	ATCC 51100	+	-
Streptococcus pseudopneumoniae	BAA-960	+	-
Streptococcus oralis	NCTC 11427	+	-
Haemophilus influenzae	ATCC 10211	+	-
Moraxella catarrhalis	ATCC 25238	+	-
Klebsiella pneumoniae	ATCC 13883	+	-
Escherichia coli	ATCC 51299	+	-
Enterococcus faecalis	ATCC 25922	+	-

[a] ATCC: American Type Culture Collection; NCTC: National Collection of Type Cultures

[b] 16S rRNA gene amplification performed as described in Woo *et al*. (Woo *et al.* 2001)

3.3.3 Evaluation of the Sequetyping Method Based on *cpsB*

The *cpsB* sequences of 138 pneumococcal strains covering 48 serotypes were tested *in vitro* (Table 3.4). These strains originate from Tanzania and the United Kingdom with different genetic backgrounds and anatomical sample sites (Table 3.4). These

strains included all of the PPV23 vaccine serotypes (VTs) except for serotype 2. In addition, 25 non-vaccine serotypes (NVTs) were tested. Altogether, the study evaluated 46 serotypes that were predicted by *in silico* analysis to generate an amplicon, and two serotypes, 27 and 29, that were not predicted to yield an amplicon.

In silico analysis based on the entire *cpsB* gene predicted that 28 of the serotypes tested would be sequetyped, and 20 serotypes tested in this study were predicted unlikely to be differentiated (Table 3.2). Of the 138 strains tested 96 (69.6%) were correctly sequetyped to the serotype. An additional 26 (18.8%) were sequetyped to the serogroup level, and 13 (9.4%) gave ambiguous results. One of 7 strains of serotype 19F (11.6402.H) was identified as a serotype 1, but at lower identity (98%) than found with the true serotype 1 strains tested. It is worth noting that the correct 19F identity was a very close second best at 97% identity (Table 3.4). At this relatively low percent identity we would express caution and assign it as ambiguous and confirm by alternative methods, as all correctly identified strains had ≥ 99% identity. In addition a 33C (SG33c) and a 36 (SG36) strain were misidentified (Table 3.4).

Table 3.4. Serotypes and strains included in this study, and sequetyping results based on *cpsB* sequences.

Serotype[a]	Sequetype	Strain	Sequence Type[b]	Country of Isolation	Isolation Source[c]	*cpsB* Accession Number
1	1	3OP2 Sep	217	Tanzania	NP	JN642309
	1	H0 8114 0126	217	UK	BC	JN660116
	1	H0 8212 0279	303	UK	BC	JN660117
	1	H0 8078 0043	618	UK	BC	JN660118
	1	H0 6274 0473	304	UK	BC	JN660119
	1	H0 8228 0507	227	UK	BC	JN660120
	1	H0 6196 0204	228	UK	BC	JN660121
	1	H0 5122 0138	306	UK	BC	JN660122
3	3	H0 9260 0327	180	UK	BC	JN660145
	3	12.1690.X	180	UK	BC	JQ743514
	3	12.1655.E	180	UK	BC	JQ743515
	3	12.1640.H	180	UK	BC	JQ743516
	3	12.1409.G	180	UK	BC	JQ743517
	3	12.1259.V	180	UK	BC	JQ743518
4	4	H0 9146 0234	246	UK	BC	JN660124
	4	H0 8102 0237	246	UK	BC	JN660125
	4	H0 7406 0041	246	UK	BC	JN664256
	4	11.6349.S	205	UK	BC	JQ743519
	4	11.5324.N	246	UK	BC	JQ743520
	4	11.4573.Q	246	UK	BC	JQ743521
5	5	H0 8034 0160	289	UK	BC	JN660126
	5	02-4520	ND	UK	BC	JQ743522
	5	04-2077	1400	UK	BC	JQ743523
	5	07-2667	289	UK	BC	JQ743524
	5	10-1351	4840	UK	Eye	JQ743525
6A	6A	H0 8212 0259	65	UK	BC	JN660127
	6A	N15	65	UK	BC	JN680106
	6A	N94	65	UK	BC	JN680118
	6A	N155	65	UK	BC	JN680128
	6A	N259	65	UK	BC	JN680138
	6A	N405	65	UK	BC	JN680146
	6A/6B	12.1624.T	396	UK	BC	JQ743526
	6A	11.6939.E	2467	UK	BC	JQ743527
	6A	11.5544.Z	1876	UK	BC	JQ743528
	6A/6B	11.5346.W	327	UK	BC	JQ743529
	6A	11.3085.J	65	UK	BC	JQ743530
6B	6B	H0 8052 0052	176	UK	BC	JN660128
	6B	H0 7156 0309	3481	UK	BC	JN660129
	6B	11NP10 Jan	4432	Tanzania	NP	JN642310
	6B	69OP10	4429	Tanzania	NP	JN642311
	6B	35NP1	4157	Tanzania	NP	JN642312
	6B	35NP6	854	Tanzania	NP	JN642313
	6B	3OP7 Sep	4373	Tanzania	NP	JN642314

Serotype[a]	Sequetype	Strain	Sequence Type[b]	Country of Origin	Isolation Source[c]	*cpsB* Accession Number
6B	6B	3NP7 Mar	4368	Tanzania	NP	JN642315
6C	6C/D	H0 5252 0052	1390	UK	BC	JN660130
7C	7C	1003-2	ND	Tanzania	NP	JN642316
	7C	2514-1	ND	Tanzania	NP	JN642317
7F	7F/7A	SG07F	ND	UK	C	JN660086
	7F/7A	12.1695.Y	191	UK	BC	JQ743531
	7F/7A	12.1675.H	191	UK	BC	JQ743532
	7F/7A	12.1666.R	191	UK	BC	JQ743533
	7F/7A	12.1426.R	191	UK	BC	JQ743534
	7F/7A	12.1299.R	191	UK	BC	JQ743535
	7F/7A	12.1276.C	191	UK	BC	JQ743536
8	8	H0 8342 0074	53	UK	BC	JN660132
	8	H0 9122 0175	53	UK	BC	JN660133
9N	9N	H0 7174 0058	66	UK	BC	JN660134
	9N	H0 7018 0063	66	UK	BC	JN660146
	9N	H0 9080 0083	66	UK	BC	JN660135
9V	9V	H0 7016 0058	156	UK	BC	JN660136
	9V	H0 8156 0265	162	UK	BC	JN660141
	9V	11.5715.C	156	UK	BC	JQ743537
	9V	11.3601.D	162	UK	BC	JQ743538
	9V	11.3021.E	162	UK	BC	JQ743539
	9V	11.2143.S	156	UK	BC	JQ743540
10A	10A	SG10a	ND	UK	C	JN660087
	10A	21NP1	ND	Tanzania	NP	JN642318
	10A	18NP10	852	Tanzania	NP	JN642319
10F	10F/10C	SG10f	ND	UK	C	JN660088
11A	11A/11D/18F	SG11a	ND	UK	C	JN660089
	11A/11D/18F	29OP10	5752	Tanzania	NP	JN642320
12B	12B	SG12b	ND	UK	C	JN660091
12F	12B	SG12f	ND	UK	C	JN660090
14	14	H0 8208 0041	9	UK	BC	JN660147
	14	H0 8396 0107	124	UK	BC	JN660142
	14	H0 8084 0056	124	UK	BC	JN660143
	14	H0 7442 0047	124	UK	BC	JN660144
	14	10-1688	ND	UK	S	JQ743541
	14	10-2893	ND	UK	S	JQ743542
15A	15A	SG15a	ND	UK	C	JN660092
15B	15B	SG15b	ND	UK	C	JN660093
16F	16F	SG16f	ND	UK	C	JN660094
17A	34/17A	SG17a	ND	UK	C	JN660095
17F	17F/33C	H0 9084 0082	392	UK	BC	JN660140
	4/17F/33C/9V	22NP2	4160	Tanzania	NP	JN642321
	17F/33C	H0 6092 0119	392	UK	BC	JN660137
	17F/33C	H0 8334 0064	964	UK	BC	JN660138
18B	18B/18C	SG18b	ND	UK	C	JN660096

Serotype[a]	Sequetype	Strain	Sequence Type[b]	Country of Origin	Isolation Source[c]	*cpsB* Accession Number
18C	18B/18C	10.2671.G	113	UK	BC	JQ743543
	18B/18C	10.2178.X	113	UK	BC	JQ743544
	18B/18C	09.2555.S	113	UK	BC	JQ743545
	18B/18C	09.2223.P	638	UK	BC	JQ743546
	18B/18C	09.1742.V	2449	UK	BC	JQ743547
	18B/18C	09.1153.K	4162	UK	BC	JQ743548
19A	19A	35NP1	4162	Tanzania	NP	JN642322
	19A	H0 9186 0354	199	UK	BC	JN664257
	19A	12.1625.M	2081	UK	BC	JQ743549
	19A	12.1623.F	461	UK	BC	JQ743550
	19A	12.1579.M	450	UK	BC	JQ743551
	19A	H0 7336 0087	276	UK	BC	JN664258
19F	19F	ATCC 49619	ND	UK	C	JN642323
	19F	16OP4	347	Tanzania	NP	JN642324
	19F	35NP8	6170	Tanzania	BC	JN642325
	19F	H0 8112 0101	162	UK	BC	JN664259
	19F	H0 8242 0108	162	UK	BC	JN664260
	19F	11.6554.S	ND	UK	Sputum	JQ743552
	1[e]	11.6402.H	ND	UK	Sputum	JQ743553
20	20/13	SG20	ND	UK	C	JN660097
21	21	45OP9 Jan	1145	Tanzania	NP	JN642326
	21	7NP4	ND	Tanzania	NP	JQ009436
22A	22F/22A	SG22a	ND	UK	C	JN660098
22F	22F/22A	SG22f	ND	UK	C	JN660099
23B	23B	SG23b	ND	UK	C	JN660100
	23B	7OP9	ND	Tanzania	NP	JN642327
23F	23F	SG23f	ND	UK	C	JN664261
	23F	11.4091.H	ND	UK	BC	JQ743554
	23F	11.3056.E	33	UK	BC	JQ743555
	23F	11.2827.S	1682	UK	BC	JQ743556
	23F	11.2737.T	1682	UK	BC	JQ743557
	23F	11.1373.Z	6959	UK	BC	JQ743558
24B	24B	SG24b	ND	UK	C	JN660101
24F	24B	SG24f	ND	UK	C	JN660102
27	27	H0 8432 0293	1475	UK	BC	JN660139
28A	28A	SG28a	ND	UK	C	JN660103
28F	28A	SG28f	ND	UK	C	JN660104
29	29	SG29	ND	UK	C	JN660105
31	31	SG31	ND	UK	C	JN660106
33B	33B	SG33b	ND	UK	C	JN660108
33C[b]	35B/35C	SG33c	ND	UK	C	JN660109
33D	33B	SG33d	ND	UK	C	JN660110
33F	33A/33F/35A	SG33f	ND	UK	C	JN660111
	33A/33F/35A	SG33fb	ND	UK	C	JN660107
34	34/17A	SG34	ND	UK	C	JN660112

Serotype[a]	Sequetype	Strain	Sequence Type[b]	Country of Origin	Isolation Source[c]	*cpsB* Accession Number
35A	35A/33F/33A	SG35a	ND	UK	C	JN642328
	35B/35C	16NP10	840	Tanzania	NP	JN660113
	35B/35C	55OP3	ND	Tanzania	NP	JQ009437
35B	35B/35C	SG35b	ND	UK	C	JN660114
36	7F/21	SG36	ND	UK	C	JN660115

[a] Serotypes covered in PPV23, PCV7, and PCV13 are in bold.
[b] ND: no data.
[c] NP: nasopharynx; BC: blood culture; C: control laboratory strain
[d] Identical: same results from serotype and sequetype; Serogroup: serogroup-specific sequetype result; Ambiguous: correct sequetype shared with another serotype from different serogroup; misidentified, sequetype gave top result of a serotype outside of the same serogroup.
[e] 11.6402 (19F) sequetyped as serotype 1, with 98% identity (715/732) with the variation at the 5' end of the gene. Note, the correct 19F serotype was identified at 97% homology.

Of the 23 VTs tested, all strains of 15 serotypes (1, 3, 4, 5, 6A, 6B, 8, 9N, 9V, 10A, 14, 15B, 19A, 19F, and 23F) were correctly sequetyped (Table 3.4), except for two out of 11 strains of 6A that were sequetyped to the serogroup level, and a strain of 19F that was ambiguous as described above (Table 3.4). An additional four VTs were sequetyped to the correct serogroup, where 7F, 12F, 18C, and 22F were sequetyped as 7F/7A, 12B, 18B/18C, and 22F/22A respectively. Ambiguous results were obtained for serotypes 11A, 17F, 20, and 33F. Except for the ambiguous 19F strain, 11.6402.H, none of the VTs tested were incorrectly identified.

Of the 25 NVTs tested, 12 were correctly sequetyped, and 8 additional NVTs were serogroup-specific. Three NVTs gave ambiguous results, in addition to the misidentified serotypes 33C and 36.

From two to 11 strains of 23 serotypes were tested (1, 3, 4, 5, 6A, 6B, 7C, 7F, 8, 9N, 9V, 10A, 11A, 14, 17F, 18C, 19A, 19F, 21, 23B, 23F, 33F, 35A), some with different multilocus sequence types (Table 3.4). All but serotypes 6A, 11A, 17F, 19F and 35A gave correct sequetype for all strains (Table 3.4). Two strains of 6A were sequetyped to serogroup, namely 6A/6B (Table 3.4). All strains of 11A gave an ambiguous sequetype of 11A/11D/18F (Table 3.4), and all but one of the serotype 17F strains gave an ambiguous sequetype of 17F/33C (Table 3.4). The remaining 17F strain also

had the top sequetype score for 4/9V/17F/33C (Table 3.4). As described above, a 19F strain was identified as serotype 1, but at a relatively low identity of 98% with 17 bp differences, that were all within the 5' end of the gene. Of the three strains of 35A that were included in this study two were from Tanzania and one was from the UK. The Tanzania strains (16NP10 and 55OP3) were both typed to the serogroup level (35B/35C), while the UK strain (SG35a) gave an ambiguous result (35A/33A/33F) (Table 3.4).

3.3.4 Comparison Between the Predicted *In Silico* and Experimental Data.

All the strains included in the study could be amplified, including serotypes 27 and 29 that were not predicted to be amplifiable (Table 3.4). These two serotypes were also sequetyped correctly.

Of the 28 serotypes in this study that were predicted to be sequetyped, 23 were correctly sequetyped. Four of the remaining serotypes (7F, 24F, 28F, and 33D) were sequetyped to the correct serogroup, and the serotype 36 strain was identified as serotype 7/21 (Table 3.4).

Twenty serotypes tested were not predicted to be sequetypable. Remarkably, two serotype 9V strains were differentiated from 9A, a serotype predicted to have an identical *cpsB*. The remaining serotypes, as predicted, were serogroup specific, consistent, or misidentified. The single serotype 6C strain in the study was sequetyped to the serotypes 6C or 6D, but not to 6A or 6B.

A breakdown of the serotypes in our study is provided in Table 3.5, according to the success of their identifications based on sequetyping.

Table 3.5 Serotypes classified according to sequetyping accuracy

Correctly Identified	Serogroup-Specific	Ambiguous	Misidentified
1, 3, 4, 5, 6A, 6B, 7C, **8, 9N, 9V, 10A,** 12B, **14**, 15A, **15B/C**, 16F, **19A**, **19F**, 21, 23B, **23F**, 24B, 27, 28A, 29, 31, 33B	**6A**[b]**,** 6C**, 7F,** 10F, **12F**, 18B, **18C**, 22A, **22F**, 24F, 28F, 33D, 35B, 35C	**11A,** 17A, **17F, 20, 33F**, 34, 35A	**19F**[b], **33C,** 36

[a] PPV23 serotypes in bold, PCV13 serotypes in red bold.
[b] Two strains of serotype 6A (12.1624.T and 11.5346.W) were serogroup specific (6A/6B), one serotype 19F strain (11.6402.H) was identified as serotype 1.

3.4 Discussion

Two primer-binding sites were located with a novel software script that was predicted to generate an amplicon in 84 of 92 serotypes sequences interrogated. Furthermore, amplification of this region using a single primer pair on pneumococcal strains isolated from different geographical regions and anatomical sites has been shown to be successful in identifying a high number of serotypes that included those in the latest conjugate vaccine formulation, as well as the current replacement serotypes associated with invasive diseases.

3.4.1 Rationale for Designing Sequetyping Method

Serotyping is conventionally based on the Quellung reaction with anti-capsular sera, but this is expensive, labor-intensive, and prone to errors. Large numbers of pneumococcal cells are needed, and false reactivity may occur with different serotypes and other streptococcal species (Browne, Miegel & Stottmeier 1984; Lee *et al.* 1984; Werno & Murdoch 2008). Immunoblotting methods have been developed and have improved sensitivity and specificity but are not used routinely in most laboratories (Bogaert *et al.* 2004b; Bronsdon *et al.* 2004).

New methods include PCR with multiple primer sets targeting serotype-specific regions of *cps* (Brito, Ramirez & De Lencastre 2003; Kong & Gilbert 2003; Lawrence *et al.* 2003; Rubin & Rizvi 2004; Kong *et al.* 2005; O'Halloran D & Cafferkey 2005; Billal *et al.* 2006; Pai, Gertz & Beall 2006), but serotype coverage is limited with the use of many primers. Microarray methods for serotyping can provide coverage of all known serotypes (Wang *et al.* 2007; Donkor *et al.* 2011; Tomita *et al.* 2011a; Turner *et al.* 2011), but its high cost would limit its routine use in resource-poor regions with high burdens of disease. A method based on serotype-specific glycosyl transferase genes was proposed (Tomita *et al.* 2011b), but only multiple strains of serogroup 6 and serotype 19F were tested. Existing methods targeting *cps* employing either multiple restriction enzymes and/or a long PCR fragment making amplification difficult and inconsistencies have been described (Lawrence *et al.* 2000; Batt *et al.* 2005).

A sequence-based typing system targeting regulatory region of *cps* had been proposed (Kong & Gilbert 2003; Kong *et al.* 2005), but a low resolution was seen between

cross-reactive serotypes, in addition to a lack of reproducibility in other serotypes. In those methodologies, differentiation of cross-reactive serotypes therefore required high number of primers to amplify the serotype specific *wzy and wzx* genes. This low resolution may have been due to the limited availability of published sequence data of only 11 serotypes at that time (Jiang, Wang & Reeves 2001).

With the recent characterization of the *cps* locus of 92 serotypes (Bentley *et al.* 2006; Park *et al.* 2007; Bratcher *et al.* 2010), and *in silico* assessment of the capsulation region, we have identified a region unique in sequence in many serotypes that is flanked by conserved primer binding sites. The region is located at the 5' end of the *cps* locus, a region previously known to be relatively conserved (Morona, Morona & Paton 1997; Morona, Morona & Paton 1999a; Jiang, Wang & Reeves 2001; McEllistrem *et al.* 2004; Bentley *et al.* 2006). PCR testing using this putative primer pair has resulted in the corrected identification of over 88% of pneumococci in our collection. It is notable that this region is only amplified in the pneumococcus in this study, with no amplification present in members of *S. mitis*, *S. oralis*, and *S. pseudopneumoniae* that have traditionally been difficult to differentiate (Arbique *et al.* 2004). However only one strain each was employed to detect specificity.

3.4.2 Sequetyping Identifies PCV13 Serotypes at Least to Serogroup

All but two of the 13 conjugate vaccine serotypes tested in our study were identified to serotype level whilst serotypes 7F and 18C were serogroup-specific. Serotypes such as 19A, 6A, and 15B/C are currently among the most prevalent replacement NVTs globally since PCV7 vaccination (Singleton *et al.* 2007; Fenoll *et al.* 2009; Huang *et al.* 2009; Hanage *et al.* 2010). These replacement types were correctly identified. Serotypes 6C, 19A and 22F (Kaye *et al.* 2009; Tocheva *et al.* 2011), which are among the common non-PCV7 serotypes causing IPD in the UK, could be identified to at least serogroup level. Reviewing previous epidemiological data demonstrated that our method should be able to correctly identify more than 76% of invasive pneumococcal diseases in the UK to at least serogroup and over 92% of penicillin-nonsusceptible clinical isolates (Cooke *et al.* 2010; Gladstone *et al.* 2011a). Thus, we conclude that this method could be used to reduce the number of strains for which conventional serotyping methods were required and make considerable cost savings.

3.4.3 *In Silico* Analysis and Experimental Results Discrepancies

Differences between the predicted *in silico* analysis data and the experimental sequetyping results were observed. Serotypes 27 and 29 that were not predicted to produce an amplicon were in fact amplified and sequetyped *in vitro*. The values inputted into the algorithm for locating appropriate PCR primer sites could be a possible explanation for this discrepancy. The potential primer-binding site mismatch tolerance was set to no more than two nucleotides, but additional primer mismatches may allow successful amplification. The threshold number of mismatches for successful amplification is not known and is likely to depend on the length and sequence of the primers. It is also possible that the locations of the mismatches within the primer also govern the success of amplification. Serotypes, 11F, 39, 43 that were predicted to be non-amplifiable were amplified in practice. As for serotypes 25A, 25F, and 38, *cps* alignment suggested that despite lacking the *cpsB* gene at the proposed location of *cps* it is possible to amplify a larger fragment (Bentley *et al.* 2006; Mavroidi *et al.* 2007). It is feasible that successful amplification of these serotypes may simply require alteration in the thermal cycling conditions without the need to employ additional primers.

Discrepancies between the predicted and experimental results where some serotypes could be sequetyped despite predictions to the contrary may be due to additional *cpsB* sequences deposited onto GenBank during this study that expanded the database available for comparison. As the *in silico* analysis was based on a representative strain of each serotype, strains of different genetic backgrounds may have variant alleles of *cpsB* (Varvio *et al.* 2009; Elberse *et al.* 2011). Indeed, this was true for serotype 9V, where sequetyping our strains matched to both serotypes 9A and 9V of the representative strains initially interrogated. However, additional 9V strains deposited in GenBank during our study matched precisely to our 9V resulting in correct identification (van Selm *et al.* 2002). Conversely, some serotypes could not be sequetyped, contrary to the *in silico* predictions. Additional strains would be required to evaluate this method more fully.

3.4.4 Reproducibility of Sequetyping Method

Multiple strains of the same serotype and different STs collected over a period of seven years could be identified correctly. All but two of the 11 serotype 6A strains tested were correctly typed, which were sequetyped to either 6A or 6B. It has been shown that serotype 6A and 6B strains of unrelated genetic backgrounds can have identical *cpsB* sequences (Elberse *et al.* 2011). A single 19F strain out of 7 could not be typed correctly, showing relatively low identity with other serotype 19F sequences on the GenBank database. It has been previously reported that intra-serotype sequence variations at the *cps* regulatory region exist (Varvio *et al.* 2009). Chapter 6 presents works comparing differences in intra-serotype variation.

3.4.5 Potential for Creation of Curated *cpsB* Database

The data presented are, in part derived from the uncurated GenBank database. The implementation of a curated *cpsB* database may potentially increase the accuracy of this method over time, by incorporating sequence data from different laboratories.

3.4.6 Identification of Large Proportion of Clinical Isolates by Sequetyping

This method is not likely to be able to replace serotyping as some cannot be identified and others only to serogroup level. Reference serotyping will be required for some isolates until further refinements are produced. The majority of the serotypes included in the latest vaccines could be identified correctly, with more identifiable at the serogroup level. This method could correctly identify some of the most common serotypes seen in IPD and carriage, including types of PCV13 and replacement serotypes (Fenoll *et al.* 2009; Huang *et al.* 2009; Sa-Leao *et al.* 2009; Jefferies, Tee & Clarke 2011). Therefore, the proposed typing method could characterize commonly-encountered serotypes, thus reducing cost and time compared to conventional serological and multiplex-PCR serotyping methods.

As the initial aim of this method is to provide a cost-effective alternative to conventional serotyping in resource-poor regions with minimal vaccine coverage, the identifiable serotypes are still likely to be common in pneumococcal disease and carriage in these regions (Leimkugel *et al.* 2005; Arifeen *et al.* 2009; Traore *et al.* 2009). Strains with ambiguous sequetype results and those without an amplicon can be assessed by conventional serological and other DNA-based methods, most likely

targeting the *wzy* and *wzx* regions (Kong *et al.* 2005). While additional *in silico* analysis is required to address the misidentified serotypes such as 33C and 36, these two serotypes are not very common and weighs little against the effectiveness of this method in identifying prevalent serotypes. The misidentified serotype 19F strains showed high identity with a serotype 1 sequence at the 5' end, and further study is needed to determine whether this was due to a recombination event. Thus. it is only with more experience of the technique that we can understand the effect of recombination events on the accuracy of the method.

3.4.7 Advantages of Sequetyping

This new sequetyping method has multiple advantages over other previously reported serotyping methods. It requires only a single PCR amplification. The reaction set up is straightforward, robust and economical requiring only crude genomic DNA from heat-lyzed cells. This creates an opportunity for routine laboratories with a PCR and sequencing facilities to serotype the majority of pneumococcal strains without the need for an expensive set of serological reagents. Excluding the cost of man-hours, sequetyping approximately costs $6.60 US per pneumococcus, more economic compared to $15 US for serotyping a pneumococcus by the Quellung reaction, as derived by Lalitha *et al.* (Lalitha *et al.* 1996). Sequetyping is likely to be a cost effective alternative to currently available microarray technologies for serotyping. More economic methods are present such as multiplex/multistep PCRs, however multiple sequential reactions, in addition to the reduced serotype coverage, are likely to be time consuming and costly when labor hours are taken into consideration. It is worth noting that the indicated sequetyping cost includes culturing on blood agar, and a culture free amplification method would reduce sequetyping time and cost. We are currently assessing the feasibility of sequetyping directly on clinical samples as well as samples known to contain multiple serotypes.

3.5 Concluding Remarks

In conclusion, our single primer sequetyping method is simple to perform, robust and economic. It is specific for the pneumococcus and can identify serotypes included in the latest conjugate vaccines and a high number of clinically relevant serotypes. It has the potential to identify new serotypes that may emerge in the post-vaccine era that are nontypable by standard methods. With the emergence of whole-genome data from high-throughout next-generation sequencing platforms, sequetyping may prove to be a long-lasting typing method by taking advantage of new sequence information, modifying its resolution power based on increased knowledge of pneumococcal *cps* sequences. We are currently evaluating our sequetyping method on pneumococcal strains of other serotypes to fully assess the coverage of sequetyping. In addition, plans are underway to make the *in silico* primer design algorithm available to other laboratory scientists.

CHAPTER FOUR: The Adaptive Potential of *Streptococcus pneumoniae* During Colonization

4.1 Introduction

4.1.1 Measurements of Pneumococcal Diversity During Colonization

Studies investigating co-colonization assess pneumococcal diversity by serotype, antibiotic susceptibility, and/or genetic diversity of co-colonizing strains.

4.1.1.1 Co-Colonization of Multiple Serotypes

Intuitively, the pre-requisite of serotype switching (Section 1.4.4) is the presence of multiple serotypes colonizing together. Due to its clinical relevance, particularly in the potential effects of serotype replacement upon use of pneumococcal vaccines, serotype is the most common marker for describing co-colonization. A child may acquire up to 6 different serotypes sequentially in the span of 2 years (Gray, Converse & Dillon 1980), and up to 5 serotypes may be present at a given time (Table 4.1). Most initial studies of multiple serotype colonization relied on culture methods, and recently numerous DNA-based methods have been employed. These methods will be discussed below.

4.1.1.2 Methodologies Used for Detecting Simultaneous Serotype Carriage

Detecting colonizations with multiple serotypes depends on the methodologies used, and insensitive methods may provide an underestimate of multiple carriage and carriage in general. The prevalence of multiple colonizations depends on the cohort studied, however the wide range in prevalence of multiple carriage may be a reflection of the different methodologies employed (Table 4.1). The World Health Organization (WHO) has devised a standardized method for detecting pneumococcal carriage (O'Brien, Nohynek & Group 2003). However, the majority of studies with extensive sampling information do not conform to the WHO protocol (Gladstone *et al.* 2011b). The earliest studies of multiple colonization involved injection of saliva samples (Gundel & Okura 1933) into the peritoneum of mice, and serotyping was performed in peritoneal samples. The presence of multiple serotypes could be detected by this method by re-inoculating the original samples into mice, however including in the sample antiserum specific for the previous serotype detected. Mouse

inoculation is sensitive for detecting multiple serotypes but it is time-consuming and costly (Huebner *et al.* 2000a).

Current culture-based studies of multiple colonizations include serotyping on numerous colonies following plating of nasopharyngeal swabs. Sensitivity of culture-based methods can be increased by inoculating swabs into an enrichment broth to detect additional colonization events and serotypes (Kaltoft *et al.* 2008). However, this may affect the relative proportion of different serotypes in the sample, and is not recommended for comparing proportions of different colonizing serotypes. Alternatively, immunoblot methods have been employed to detect multiple serotypes on agar plates with more than 25 colonies (Bronsdon *et al.* 2004; O'Brien *et al.* 2007). Molecular methods by means of PCR-amplification of serotype-specific genes provide a much more sensitive approach (Billal, 2008), however to achieve maximal serotype coverage of able to detect all 93 serotypes present, microarray may be a promising alternative (Turner *et al.* 2011). Different methods of serotyping are described in Chapter 3.

4.1.1.3 Serotyping by Colony Characterization on Agar Plates

While characterizing one colony per sample provides a representation of the prevalent serotype in the population, detecting co-colonizing and rare serotypes require serotyping additional colonies (Charalambous, Oriyo & Gillespie 2008). Detecting colonizations with multiple serotypes depends on the methodologies used; insensitive methods selecting low number of colonies or infrequent sampling (in a longitudinal study) may affect isolation rates, potentially underestimating pneumococcal carriage (Vives *et al.* 1997; Huebner *et al.* 2000a). Most culture-based studies of multiple colonizations perform serotyping on numerous colonies following plating of nasopharyngeal swabs, and additional serotypes may be identified if more colonies were characterized. Table 4.1 lists the studies of culture-based multiple colonization conducted, and their methodologies.

4.1.1.3.1 *Nasal Swab/Wash/Aspirate and Oral Swab in Revealing Serotype Diversity*

Abdullahi *et al.* (Abdullahi *et al.* 2007) compared nasal wash and swab for their sensitivity of detecting multiple pneumococcal carriage, and concluded that although wash was more sensitive in detecting pneumococcal colonization, swabs detected

more serotypes colonizing simultaneously. Higher number of serotypes within a nasopharynx can also be detected by the swab method (three serotypes by swab versus two by wash). There is no difference in swabbing the left or the right nostril. Nasal aspirate may be marginally more sensitive than nasal swabbing in detecting colonizing pneumococci (Rapola *et al.* 1997). If nasal secretions are unavailable, nasal swab is more sensitive than oral swabs, but sampling from the oropharynx may reveal additional colonization events (Masters *et al.* 1958; Gray, Converse & Dillon 1980; Capeding *et al.* 1995; Watt *et al.* 2004). Charalambous *et al.* (Charalambous *et al.* 2008) postulate that serotyping pneumococci from both the nose and throat would provide more accurate representation of multiple colonization than either site alone.

4.1.1.3.2 *Colony Serotyping Based on Morphology or Random Selection*

Hare *et al.* (Hare *et al.* 2008) found that selecting colonies randomly for serotyping increases likelihood of detecting multiple serotypes compared to serotyping colonies based on colony morphology (17% multiple colonization prevalence by random selection and 14% by morphology). However, the study also addressed that characterization of colonies based on random selection require more colonies to achieve comparable rate of multiple carriage as measured by morphology.

4.1.1.3.3 *Enrichment Broth*

Serotyping directly from enrichment broth can increase the yield of rarely detected serotypes (Kaltoft *et al.* 2008; Vestrheim *et al.* 2008a). However, extended enrichment incubation time (<6 hours) may result in overgrowth of predominant serotypes and masking of other serotypes in the sample (Antonio *et al.* 2009; Carvalho *et al.* 2010). Masking by predominant serotypes would increase the difficulty in detecting rarer serotypes, hence provide an underestimate of multiple carriage.

4.1.3.3.4 *Binomial Formula in Assessing Sensitivity of Serotyping*

The binomial formula described by Huebner (Huebner *et al.* 2000a), derived as $q^n = (1-P)$, where q = 1 - proportion of organisms in co-colonization, n = number of colonies characterized, and P = confidence interval, provides a mathematical relationship between the number of colonies analyzed and the probability of detecting a serotype of a particular prevalence. This formula calculates the amount of colonies

required for a serotype of a particular prevalence to be detected with a given confidence. For example, to detect serotypes with a proportion of 5% in the nasopharynx with 95% confidence, fifty-nine colonies would need to be examined. Nearly 300 colonies would be required for serotypes with proportion of 1%. Therefore, detecting multiple colonization by selecting colonies may provide an underestimate of the serotype diversity within a colonization event unless a large number of colonies are selected. Comparison of different methods for detecting multiple carriage revealed that serotyping a colony sweep of pneumococci provides co-colonization rates comparable to that done with DNA-based methods (Turner *et al.* 2011).

4.1.1.3.5 *Summary for Culture Methods for Serotyping*

Multiple colonization assessments by means of culturing have been employed commonly, however current methods vary in terms of isolation sources and methods, number of colonies selected, and the selection criteria for colony characterization. Therefore, comparison of co-colonization prevalence between different studies may prove difficult. Also, rates of multiple colonization are also likely to be dependent on the pneumococcal prevalence of cohort being studied. Risk factors such as young age, history of and exposure to smoking, immunosuppression, and other factors likely to contribute to increased pneumococcal colonization and transmission may also contribute to multiple colonization, however these factors have not been directly linked to increased likelihood for multiple colonization.

Table 4.1. Selected studies of culture-based serotyping methods for detecting multiple colonizations

Year	Country	Age group	Detection Method[a]	Multiple Pneumococcal Prevalence (%)[b]	Max No. Serotypes Detected	Reference
1933	Germany	12-14 yrs	Intra-peritoneal injection into mice	74	4	(Gundel & Okura 1933)
1971	US	Families	Pharyngeal swab, colony number not specified	1.3 (of pos)	2	(Dowling, Sheehe & Feldman 1971)
1985	Australia	Children+Adult	Up to 4 colonies per NP swab	5	3	(Hansman *et al.* 1985).
1989	PNG	2 wks–12 yrs	Up to 50 colonies per nasal aspirate, six per swab	29.5 (of pos)	3	(Gratten *et al.* 1989).
1990	PNG	<5 yrs	Four colonies by morphology	33	4	(Montgomery *et al.* 1990)
1994	Australia	< 13 yrs	Four colonies per aspirate by morphology or randomly	21 (of pos)	3	(Gratten *et al.* 1994)
1996	The Gambia	< 2 yrs	Unknown	23	Unknown	(Obaro *et al.* 1996)
1997	Finland	<7 years	CIE and LA on NP/OP swabs/aspirates	16	2	(Rapola *et al.* 1997)
2000	SA	1-60 mts	3-5 colonies per NP swab	1.3	2	(Huebner *et al.* 2000a)
2000	Israel	1-60 mts	Three colonies per swab	2.4	2	(Huebner *et al.* 2000a)
2002	Portugal	6-72 mts	6-8 colonies per swab	10.8	2	(Sa-Leao *et al.* 2002)
2003	UK	1-24 mts	Three colonies per swab	3.3[c]	3	(Meats *et al.* 2003)
2004	US	< 5 years	Immunoblot on agar plate + four colonies per sample	19	2	(Bronsdon *et al.* 2004)
2005	UK	Children+Adult	Two colonies by morphology	0.16	Unknown	(Hussain *et al.* 2005)
2007	US	<2 years	Immunoblot on agar plate	8.1 (of pos)	3	(O'Brien *et al.* 2007)
2008	Australia	Children	Four colonies, randomly picked or colony morphology	20	3	(Hare *et al.* 2008)
2008	Denmark	Mean 23-54 mts	Enrichment broth	9.9 (of pos)	3	(Kaltoft *et al.* 2008)
2008	Kenya	< 86 yrs	Four colonies based on morphology	3.7	2	(Abdullahi *et al.* 2008)
2008	Norway	1-5 yrs	Enrichment broth and by plating of up to 16 colonies	9.4	3	(Vestrheim *et al.* 2008a)
2009	Denmark	< 3 years	Enrichment broth from swab	10.9 (of pos)	3	(Auranen *et al.* 2010)
2011	Tanzania	1-5 years	Throat + nose swabs, up to 20 colonies total	28.5 (of pos)[e]	5	(Leung *et al.* 2011)
2011	Thailand	< 3 years	Two colonies	11.2	2	(Turner *et al.* 2011)
			Sweep colony serotyping	43.2	4	(Turner *et al.* 2011)

[a] PNG: Papua New Guinea, SA: South Africa, mts: Months, CIE: Counterimmunoelectrophoresis, LA: Latex Agglutination, NP: Nasopharyngeal, OP: Oropharyngeal

[b] Prevalence calculated as percentage of cohort, unless indicated by "of pos" meaning of pneumococcal-positive subjects

[c] Rate of co-colonization was not explicitly stated in literature. Text describes that multiple serotypes were found in 80.4% (74/92) of children over a two-year period. Presence of simultaneous colonization was seen in Table 4 of literature where only samples with more than six serotypes over two years were shown. The percentage of 3.3% is the number of colonization events with multiple serotypes simultaneously from this table divided by the total number of samples, and therefore 3.3% may be an underestimate of the true percentage of simultaneous colonization.

[d] Nasopharyngeal samples taken from children with lower airway infection.

[e] Works included in this chapter. Co-colonization prevalence calculated as percent of all pneumococcal-positive colonization events rather than swab or individual

4.1.1.4 DNA-based Methods for Detecting Serotype Diversity

DNA-based serotyping methods including multiplex PCR, real-time PCR, and microarray have been applied to assess co-colonizing serotypes (Table 4.2). Superiority in sensitivity with DNA-based serotyping methods compared to culture methods has been demonstrated (Azzari *et al.* 2008; Tarragó *et al.* 2008; Carvalho *et al.* 2010). Importantly, the increased sensitivity indicates that detection of rare serotypes is made more efficient, as these less-abundant serotypes may play crucial roles in the pneumococcal epidemiology following vaccine interventions (see below). These methods have also revealed multiple serotypes in cases where conventional cultures could not (Billal *et al.* 2008; Carvalho *et al.* 2010). With microarray testing of multiple colonization, serotype-specific regions from nine different serotypes could be detected in a colonization sample, compared to two serotypes if the same sample was performed by serotyping two colonies on blood agar plate (Turner *et al.* 2011). Given the costs of antisera, multiplex PCR is a cost-effective alternative to serological methods, important in resource-poor laboratory settings with high pneumococcal carriage prevalence.

Table 4.2. Selected studies of DNA-based serotyping methods for detecting multiple colonizations

Year	Country	Age group	Method[a]	Number (+ Serogroups targeted)	Multiple Pneumococcal Prevalence (%)[b]	Maximum No. Serotypes	Reference
2005	Columbia	< 5 yrs	mPCR	7 (+ groups 6 and 18)	9%	2	(Moreno *et al.* 2005)
2008	Japan	1-5 yrs	mPCR	7 (+ groups 6, 18, 19, 23)	6%	2	(Billal *et al.* 2008)
2009	Gambia	Children+Adult	mPCR	28 (+ 6A/B)	32%[c]	3	(Antonio *et al.* 2009)
2009	Venezuela	< 7 yrs	mPCR	7 (PCV7 serotypes)	20%	2	(Rivera-Olivero *et al.* 2009)
2010	Italy	unknown	mPCR	31	16%	2	(Azzari *et al.* 2010)
			Real-Time PCR	31	63%	4	(Azzari *et al.* 2010)
2010	Brazil	< 3 yrs	Real-Time PCR	9 + 7F/A + 9V/A + 10F/10C/33C + serogroup 24)	12%	2	(Carvalho *et al.* 2010)
2011	Switzerland	Unknown	mPCR+Microarray	91	3.4%/3.7%[d]	3[e]	(Brugger *et al.* 2010)
2011	Gambia	< 2 yrs	Microarray	91	19%	3	(Donkor *et al.* 2011)
2011	Thailand	< 3 yrs	Microarray	91	49%	9	(Turner *et al.* 2011)
2012	Valente	18-71 mts	mPCR+Microarray	91	8%/18%[d]	6	(Valente *et al.* 2012)

[a] mPCR: Multiplex-PCR
[b] Prevalence calculated as a percentage of cohort, rather than of pneumococcal-positive samples, unless indicated by "of pos" meaning of pneumococcal-positive subjects
[c] Prevalence limited to colonizations detected by the 29 serotypes covered in the study
[d] Pre-PCV7 Era % / Post-PCV7 Era
[e] Three distinct strains as detected by *ply*PCR method, maximum number of serotypes detected unknown

4.1.2 Antibiotic Susceptibility Diversity

Simultaneous colonization of different antibiotic susceptibilities allows for the transfer of resistance determinants between pneumococci. Pneumococci that are highly competent and hyper-recombinant (based on evidence of mosaic genes in concatenated multilocus sequence types) are significantly associated with resistance to antimicrobials (Hanage *et al.* 2009), and associations have been demonstrated between commonly colonizing serotypes, frequent recombination, and penicillin resistance (Hsieh *et al.* 2006). Thus, serotypes 6B, 14, and 19F (commonly encountered serotypes) were reported to have higher transformation efficiencies following competence activation *in vitro*, and showed higher prevalence of penicillin-nonsusceptibility. In contrast, serotypes 3 and 18C had the lowest transformation efficiencies among the serotypes tested, and all of the pneumococci of these serotypes were penicillin susceptible.

A pilot study in Portugal revealed co-colonization with strains of different susceptibilities to penicillin, tetracycline, and clindamycin (Sa-Leao *et al.* 2002). The simultaneous colonization of penicillin-susceptible and non-susceptible pneumococci has also been reported previously, with two strains of serotype 6B showing a 50-fold difference in penicillin minimum inhibitory concentrations (MICs) (Gratten *et al.* 1989). Tetracycline and clindamycin nonsusceptibilities can be attributed to acquisition of *tet*(M) and *erm*(TR) genes, which can be flanked by mobile elements subjected to recombinational events (Camilli *et al.* 2008; Croucher *et al.* 2011). Penicillin resistance has been associated with mosaic *pbp* genes that are acquired horizontally (Section 1.3.2). Co-colonization of strains with different susceptibilities therefore potentially facilitates the horizontal spread of drug-nonsusceptible genetic determinants.

4.1.3 DNA-based methods and MLST for detecting Pneumococcal Diversity

Indexing genetic diversity within a co-colonization event has been determined by using restriction digest of either a single gene region (*ply*NCR) (Brugger, Hathaway & Mühlemann 2009; Brugger *et al.* 2010) or whole genome (pulsed-field gel electrophoresis, PFGE, (St. Sauver *et al.* 2000), or sequence analysis of housekeeping genes (MLST). MLST in the pneumococcus involves the sequence analysis of seven housekeeping gene fragments subjected to neutral selective pressures, and the

combination of alleles of these seven loci determines the sequence type (ST) of the strain, which is a representation of the strain's genotype. Indexing variation using nucleotide sequence is unambiguous and results can be compared between laboratories. MLST is also suitable for typing of transformable species, as a single nucleotide change is equally different from a recombination of multiple nucleotide alterations. MLST was originally designed for typing within strains of another recombinant organism *N. meningitidis* (Maiden *et al.* 1998), and was soon designed for pneumococcal typing (Enright & Spratt 1998). Characterization of STs has shown that strains of the same serotype have different STs (Enright & Spratt 1998; Brueggemann *et al.* 2003; Meats *et al.* 2003; Inverarity *et al.* 2010). These observations highlight the possibility of underestimating the true diversity of pneumococcal colonization by analyzing serotype alone.

4.1.4 Representation of MLST Data: eBURST

The relationships between strains can be represented by MLST data that can be visualized as an eBURST representation of genetic relationship between strains, grouping strains sharing 6 of 7 loci into a clonal complex (CC) of related strains (Feil *et al.* 2004). Illustrating genetic relationship between strains by eBURST is advantageous over the more traditional methods such as UPGMA when large samples are employed in population studies. More importantly, it does not assume evolutionary relationship and directionality between strains, which is important in an organism such as the pneumococci, where frequent recombination events may disrupt evolutionary signals.

4.1.5 Aim

The aim of works covered in this chapter is to investigate the diversity of colonizing pneumococci in healthy Tanzanian children. Two critical phenotypes, serotype and antibiotic susceptibility to penicillin and co-trimoxazole, would be determined for these pneumococcal strains. In addition, genotyping by MLST would be performed to determine the genetic diversity of numerous colonization events, including events where co-colonization of pneumococci of multiple phenotypes was present.

4.2 Materials and Methods

Please refer to Appendix Table A1 for list of strains included in this chapter. Please refer to Chapter 2 for sample collection (Section 2.2), DNA extraction (Section 2.8.1), and standard PCR master mix components (Section 2.8.2). Gel electrophoresis, DNA purification, cycle sequencing, sequence reaction purification, and sequence analysis steps were performed as described in Sections 2.8.3 to 2.8.8. Antibiotic susceptibility and minimum inhibitory concentrations (MICs) for each strain was performed as described in Section 2.7. Antibiotypes were defined as different if they had a minimum inhibitory concentration (MIC) that varied by at least four-fold. MIC determinations were performed on three independent clones of strains within an individual colonization event with different antibiotypes.

4.2.1 Multilocus Sequence Typing (MLST)

Fragments of seven housekeeping genes were selected as described in the work of Enright *et al.* (Enright & Spratt 1998). PCR thermal cycle condition for MLST was an initial denaturation step at 95°C for 4 min, followed by 30 cycles of a 95°C denaturation step for 30 sec, a 50°C re-annealing step for 30 sec, and an extension step at 72°C for 30 sec. Consensus sequences were trimmed to length of respective locus fragments as indicated at the MLST database (http://www.mlst.net). Sequences were interrogated on the MLST database, and any novel MLST locus alleles (alleles that were previously not present in the database) were submitted to Miss Cynthia Bishop (Imperial College London) for curating and chromatogram quality control. In strains where the combination of MLST locus alleles were not observed before, the allele integers in the order of *aroE-gdh-gki-recP-spi-xpt-ddl* were also submitted to the curator of MLST database, and a novel ST was allocated to these strains.

4.2.2 eBURST

Genetically related strains are defined as having either six out of seven identical loci conserved (single locus variant; SLV), or five out of seven identical loci conserved (double locus variant; DLV) (Feil *et al.* 2004). SLV and DLV pairs were subsequently verified by repeating the MLST analysis on at least three independent clones from the frozen archives. The chromatograms were independently verified by Dr. Clare Ling.

4.3 Results

4.3.1 General Carriage Epidemiology

The cohort consisted of 83 healthy children under the age of 6 (mean age 28.7 ± 15.1 months, SD) residing in an isolated community on a sugar plantation recruited from 20th January to 7th February, 2003. Children on antibiotics within 3 months prior to recruitment were excluded from the study. Of eighty-three children recruited to the study, twenty-one children with questionnaire information were colonized by *S. pneumoniae* (carriage rate 25%) and these were followed up at monthly intervals where possible. Due to difficulty in follow-up sampling (result of discomfort following nasopharyngeal swabbing, children absent on the day of swabbing), on average 14 ± 4.4 (67%) of the children were swabbed per month. Each child was swabbed on average 7.0 ± 2.6 occasions and was positive for pneumococci on 3.0 ± 1.2 occasions. A total of 61 pneumococcal colonization episodes were captured. An average of 6.7 ± 1.7 (SD) isolates were sub-cultured and characterized per colonization event. There had been exposure to antibiotics in the previous month to sampling in 44.3% (27/61) of the pneumococcal colonization events.

Of the 61 pneumococcal colonization events in 21 children, seven (11.5%) colonization events from 6 children had multiple phenotypes (Table 4.3). All seven colonization events had multiple serotypes, and four of these also had multiple antibiotypes. All the multiple antibiotypes had penicillin MICs greater than 4 times between strains and two also had multiple MICs for Sxt. One child (Child 10) experienced two co-colonization events.

Of the 61 pneumococcal colonization events, fifteen were genotyped by multi-locus sequence typing (MLST). Five of these colonization events had multiple serotypes and antibiotypes, and 10 colonization events were single phenotypes. In total 120 isolates were phenotyped, and sequence types were determined for 98 (81.7%) of them.

4.3.2 Nasopharyngeal Colonization by Multiple Serotypes

A total of 13 serotypes were detected in this study including 1, 4, 6B, 10A, 11A/D, 13, 17F, 18B, 19A, 19F, 21, 34, and 35A. A nontypable strain was captured from a

single colonization event, based on the failure to agglutinate with pooled antisera and the absence of the *cpsA* and *cpsB* genes. Of the 7 colonization events with multiple serotypes, five colonization events had 2 serotypes, one event had 3, and up to 5 serotypes (6B, 10A, 19A, 19F, 21) were detected in Child 35 in January (Table 4.3). All events were detected in females. The most common serotypes detected within the 7 multiple phenotype colonization events was 6B (3/7, 33%) followed by 19F (2/7, 29%) and 35A (2/7, 29%). We investigated the proportions of each serotype co-colonizing and found that a dominant (as defined in this study as ≥ 80% prevalence within a colonization) serotype was found in 3 of the 7 (42.9%) colonization events, all of which had two serotypes each (Table 4.3). A dominating serotype was not observed for colonization events with more than two serotypes, but serotype 35A made up ≤ 10% of the population characterized in Child 16, which carried with 3 serotypes.

4.3.3 Nasopharyngeal Colonization by Multiple Antibiotypes

All strains were susceptible to tetracycline, chloramphenicol, erythromycin, and amoxicillin, while some strains were non-susceptible to penicillin and co-trimoxazole by disc diffusion. We determined the antibiotypes of strains, measured by MICs of penicillin and co-trimoxazole, by E-test. Multiple antibiotypes were detected in 4 colonization events, all of which also had multiple serotypes. All of the 4 events differed in penicillin MICs, with a mix of penicillin-susceptible and intermediate-resistant strains. In addition, two of the four also had multiple Sxt antibiotypes, with Child 1 colonized simultaneously by Sxt-susceptible and resistant strains, and Child 3 in September co-colonized by intermediate and resistant strains. Each of these 4 colonizations had 2 different antibiotypes. Differences in antibiotypes are associated with differences in serotypes; none of the colonizations had a different antibiotype within the same serotype. Across different colonization events studied, serotype 6B was associated with multiple antibiotypes, differing in both penicillin and Sxt (Table 4.3, Child 11 January and Child 3 March). In Child 29, serotypes 35A and 6A were detected with the same antibiotype for both penicillin and Sxt. In children where multiple colonization events were analyzed a number of times (Child 3 and Child 11), strains of the same serotypes detected in each child did not change in antibiotypes over time.

Table 4.3. Serotypes, MICs, and sequence types of pneumococci isolated from healthy Tanzanian children

Child#	Gender	Baseline Age (Months)	Colonization (Sampling Month)	No. Isolates Characterized	Serotype[a]	MIC[b]		Sequence Type[c]	Prevalence (%)
						Penicillin	Sxt		
1	F	18	January/February	6	NT	0.032	0.19	4156	5 (83.3)
					6B	0.19	8	4429	1 (16.7)
3	F	41	March	3	6B	0.032	0.125	4368	3 (100)
3	F	41	July	5	1	<0.016	<0.002	217	5 (100)
3	F	41	September	8	1	<0.016	<0.002	217	5 (62.5)
					6B	0.25	6	4373	3 (37.5)
10	F	60	January/February	7	13	0.125	3	ND	1 (14.3)
					19A	0.125	6	ND	6 (85.7)
10	F	60	March	6	13	0.125	6	4370	6 (100)
10	F	60	April	7	18B	<0.016	<0.002	ND	6 (85.7)
					6B	<0.016	<0.002	ND	1 (14.3)
11	F	27	January/February	15	6B	0.19	6	4432	15 (100)
11	F	27	May	8	4	<0.016	<0.002	ND	8 (100)
11	F	27	June	8	6B	0.19	4	4432	8 (100)
16	F	59	April	11	19F	0.19	6	347	6 (54.5)
					34	0.032	4	4158	4 (36.4)
					35A	0.25	4	840	1 (9.1)

18	F	21	January/February	3	10A	0.125	4	852	3 (100)
22	M	60	January/February	5	17F	0.016	0.125	4160	5 (100)
24	F	25	October	3	19A	0.25	6	4162	3 (100)
29	F	22	November	2	35A	0.25	4	840	1 (50)
					11A/D	0.38	6	5752	1 (50)
35	F	27	January/February	7	6B	0.125	6	854	1 (14.3)
						0.125	6	4157	1 (14.3)
					10A	0.38	6	852	2 (28.6)
					19A	0.125	2	4162	1 (14.3)
					19F	0.19	2	6170	1 (14.3)
					21	0.064	3	1145	1 (14.3)
45	F	47	January/February	9	21	0.094	4	1145	9 (100)
69	M	12	January/February	7	6B	0.19	8	4429	7 (100)

[a] NT: strains were nontypable by pooled sera with no amplification from PCR targeting *cpsA* (Pai, 2006) and *cpsB* (Leung, accepted)
[b] MIC: minimum inhibitory concentrations indicated as the median of triplicate results
[c] ND: no data
[d] Events with multiple serotypes in red, multiple antibiotypes in blue for penicillin and orange for co-trimoxazole.

4.3.4 Nasopharyngeal Colonization by Multiple STs

ST was used as a representation of the strain genotype on the 15 colonization events analyzed. A total of 18 STs were identified, of which only 7 (ST217, 347, 840, 852, 854, 1145, and 5752) were STs previously deposited on the MLST database. Searches in the MLST database revealed that all of these previously deposited STs included strains isolated in Africa (Appendix Table A3). Multiple STs were observed on the five multiple phenotype colonization events sequence typed. Up to 6 STs (852, 854, 1145, 4157, 4162, and 6170) were detected within a single colonization event from Child 35. This event had 5 serotypes (6B, 10A, 19A, 19F, and 21). STs 840, 852, 1145, and 4162 were found in more than one child where each ST expressed the same phenotype (Table 4.3). ST217 was found in Child 3 two months apart (Table 4.3). ST4432 was colonizing in Child 11 in Jan/Feb as well as in June (Table 4.3). We found multiple STs expressing the same serotype (Table 4.4). Two STs expressing serotype 6B was found in Child 35 concomitantly (Table 4.3).

Table 4.4. Serotypes with multiple sequence types

Serotype	aroE	gdh	gki	recP	spi	xpt	ddl	**ST**	**Child**	**Colonization Event**
6B	51	66	230	1	6	1	8	4432	11	Jan/Feb
	51	66	230	1	6	1	8	4432	11	June
	7	9	1	10	9	1	14	4368	3	March
	1	5	54	38	15	288	15	4373	3	September
	51	66	1	1	6	1	108	854	35	Jan/Feb
	7	9	1	63	9	1	14	4157	35	Jan/Feb
	51	66	230	1	6	1	6	4429	1	Jan/Feb
	51	66	230	1	6	1	6	4429	69	Jan/Feb
19F	12	8	9	3	3	20	57	347	16	April
	12	8	9	3	6	20	57	6170	35	Jan/Feb

4.3.5 Genetic Relatedness of STs in Cohort

eBURST was performed to determine the genetic relatedness of the STs characterized in our study. eBURST is based on a clonal expansion and diversification model that provides a method for inferring clonal relationships from MLST data (Feil *et al.* 2004). Genetically-related strains were defined as single (SLV) or double locus variants (DLV). All but 7 STs were singletons. Three pairs of SLVs were identified in different colonization events from different children (Tables 4.3 and Tables 4.5). ST6170 and ST347 (both serotype 19F and the same antibiotype), differed in the *spi*

locus by a single nucleotide (Table 4.5). ST4432 and 4429 (both serotype 6B and the same antibiotype) differed in *ddl* locus by 4 nucleotides (Table 4.5). ST4157 and ST4368 (both 6B, but with a different antibiotype for Sxt) differed in the *recP* locus by 3 nucleotides (Tables 4.3 and 4.5).

In the eBURST analysis algorithm double-locus variants (DLVs), as defined by a difference of 2 MLST loci out of 7, are only linked to a CC if an intermediate SLV is also present in the sample. However, DLVs could belong to the same CC even if the SLV intermediate was undetected in our cohort. The eBURST algorithm can also identify DLVs. ST854 (serotype 19F) is a DLV of ST4432 (serotype 6B) and ST4429 (6B), and both vary at the *ddl* and *gki* loci by more than 40 nucleotides (Table 4.5). All three STs have the same antibiotype (Table 4.3). Allele 108 has 95% identity to the *ddl* fragment of *S. mitis* (accession numbers EU075717 and EU075693) and *S. oralis* (accession number AJ387988.1).

Table 4.5. Allelic variation between single and double locus variants.

Serotype	SLV				
	Sequence Type A	Sequence Type B	Variant Locus	Variant Alleles (A/B)	Nucleotide Substitutions (% identity)
19F	347	6170	*spi*	3/6	1 (99)
6B	4432	4429	*ddl*	8/6	4 (99)
6B	4157	4368	*recP*	63/10	3 (99)
	DLV				
6B	854	4432	*ddl*	108/8	38 (92)
			gki	1/230	5 (99)
6B	854	4429	*ddl*	108/6	40 (91)
			gki	1/230	5 (99)

4.3.6 Novel STs and HGT Between Co-Colonizing Strains

A total of 11 novel STs were detected in this cohort. Some of these STs form part of a CC of SLVs found exclusively in Africa, as shown in Appendix Table A3. Given the high number of new STs detected in our cohort, we hypothesized that co-colonizing strains act as a reservoir for the assortment of housekeeping gene fragments, thus creating novel combinations of MLST loci. From a total of 11 novel STs uncovered in this study, eight of them were found co-colonizing with other strains. We compared

these eight STs to the MLST database to see if our novel STs were SLVs/DLVs of pre-existing strains, and whether the varying allele(s) are present among co-colonizing strains (i.e. a variant locus may have been acquired from a co-colonizing strain, giving rise to an SLV). Three novel STs in our cohort (ST4162, ST4157, ST6170) are related to pre-existing STs. All three novel STs were detected within Child 35. Comparison of pre-existing, novel and co-colonizing STs shows that HGT of MLST loci from co-colonizing strains in this child may have likely given rise to the new STs (Table 4.6). While in some cases the variant alleles between the SLVs differ by only one or two nucleotides (hence the variation is probably explained by mutation), it is highly improbable that mutational events alone gave rise to the difference between loci with up to 14 nucleotides. Therefore, our results suggest that novel STs such as ST4157, ST6170, and ST5752 have possibly acquired variant alleles from co-colonizing strains through genetic exchange.

Table 4.6. Comparison of housekeeping loci of novel, pre-existing, and co-colonizing sequence types within Child 35.

Case	ST	*aroE*	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*	serotype
1	854 (co-colonizing)	51	66	1	1	**6**	1	108	6B
	6170 (co-colonizing)	12	8	9	3	**6**	20	57	19F
	4162 (novel)	7	11	74	1	**6**	112	64	19A
	5270 (existing)	7	11	74	1	**17**	112	64	19A
						spi alleles 6 and 17 differ by 2 nucleotides			
2	852 (co-colonizing)	16	**9**	75	3	**9**	115	5	10A
	4157 (novel)	7	**9**	1	63	**9**	1	14	6B
	6041 (existing)	7	**62**	1	63	**6**	1	14	6C
		gdh alleles 9 and 62 differ by 3 nucleotides; *spi* allele 9 and 6 differ by 14 nucleotides							
3	854 (co-colonizing)	51	66	1	1	**6**	1	108	6B
	6170 (novel)	12	8	9	3	**6**	20	57	19F
	347 (existing)	12	8	9	3	**3**	20	57	19F
						spi alleles 6 and 3 differ by 1 nucleotide			
4	854 (co-colonizing)	51	66	1	**1**	**6**	1	108	6B
	852 (co-colonizing)	16	9	75	**3**	**9**	115	5	10A
	4162 (co-colonizing)	7	11	74	**1**	**6**	112	64	19A
	6170 (novel)	12	8	9	**3**	**6**	20	57	19F
	5769 (existing)	12	8	9	**1**	**3**	20	57	19F
					recP alleles 3 and 1 differ by 3 nucleotides; *spi* allele 6 and 3 differ by 1 nucleotide				

4.4 Discussion

We have determined the pneumococcal strain diversity within a colonized host by characterizing a random selection of up to 20 colonies from pneumococcus-positive nasopharyngeal samples. Two critical phenotypes, the serotype and antibiotic susceptibility as well as ST, were characterized. Our data revealed the extensive and so far unrecognized potential for genetic exchange in the nasopharynx of children in Tanzania. Nineteen strains were detected in 12% of the colonization episodes with extensive serotype, antibiotype and ST heterogeneity within a colonization event that had the potential to utilize the supragenome. This included the presence of antibiotic-susceptible and non-susceptible strains colonizing together. These data are important as there are few detailed studies addressing multiple pneumococcal carriage. To our knowledge, only a handful of reports in Europe (Sa-Leao *et al.* 2002; Vestrheim *et al.* 2008a) included serotyping, antibiotyping and genotyping data of multiple colonizations.

4.4.1 Cohort Characteristics

This study was conducted on children under the age of 6. Studies from Australia (Hansman *et al.* 1985) and the UK (Hussain *et al.* 2005) showed that co-colonization is more common in children than in adults. Children may have higher rates of co-colonization mainly because of the higher general colonization seen in this age group, facilitating higher rates of transmission between individuals. This cohort of children represents a semi-isolated community residing on a sugar plantation without a pneumococcal vaccination programme. Thus, our observations provide an insight into the diversity of pneumococcal colonization in the absence of pneumococcal immunization. Two recent studies have analyzed the effects of pneumococcal vaccination on the carriage of multiple serotypes (Brugger *et al.* 2010; Valente *et al.* 2012). In Switzerland, the prevalence of multiple serotype carriage did not alter three years after the use of PCV7 (Brugger *et al.* 2010). In Portugal, where PCV7 coverage reaches at least 70%, vaccinated individuals showed significantly lower rates of co-colonization than non-immunized individuals and that of the pre-vaccine era (Valente *et al.* 2012). These differences may be due to cohort differences (age, healthy or infected hosts) and methodologies (with or without culture) used for detecting carriage. Conjugate vaccination in developing countries is now being introduced (Global Alliance for Vaccines and Immunizations 2012). Thus, should this area be

subjected to future vaccination schemes, this data would be critical baseline data for studying the impact of vaccination on the ecology of multiple carriage in this region.

4.4.2 Current Work in Relation to Other Co-Colonization Studies

Many multiple colony carriage studies were limited to serotyping by serological methods and used different culture methods (Hansman *et al.* 1985; Huebner *et al.* 2000a; Meats *et al.* 2003; Bronsdon *et al.* 2004; O'Brien *et al.* 2007; Hare *et al.* 2008; Hill *et al.* 2008; Kaltoft *et al.* 2008; Auranen *et al.* 2010). These studies showed large variation of co-colonization rates (1.3% to >30%). In this study, we have employed random colony selection as it has been shown to reveal higher rate of multiple colonization compared to colony selection by morphology (Hare *et al.* 2008). However, randomization in our study referred to a selection of ten colonies per plate, instead of picking six colonies and characterizing four of them determined by a random number table as described by Hare *et al.* (Hare 2008). Therefore, a subjective experimental bias may be generated. A more systematic and objective method could be performed by drawing a 10x10 grid on the underside of the plate, and picking a single colony from each column or row of the grid.

In this study, the proportion of episodes with multiple serotypes were detected to be within the range published by others, and we have detected up to five serotypes colonizing together. We have identified the co-colonization of different strains of a serotype, similar to a finding made previously in Bolivian school children (Inverarity *et al.* 2010). This supports that serotyping alone underestimates the diversity of the colonizing population of an organism.

Based on the binomial formula (Huebner *et al.* 2000a), our method would reveal serotypes that constitute 14% of the nasopharynx with 95% confidence. However, a correlation between the number of colonies characterized and the prevalence of multiple strain colonizations cannot be made from these studies as different methodologies were used. The cohort studied, colonizing serotypes and their durations of carriage are also likely to influence the prevalence of co-colonization and the extent of serotype diversity. DNA-based methods by means of PCR-amplification and microarrays will enhance sensitivity with optimal serotype detection coverage (Billal *et al.* 2008; Donkor *et al.* 2011; Turner *et al.* 2011). Serotyping by colony

sweeps may detect comparable levels of multiple colonization as microarray, but colony sweep serotyping is problematic in detecting rare colonizers due to the inferior sensitivity (Turner *et al.* 2011). Microarray is also limited to a handful of reference laboratories due to its high cost, and is unlikely to be widely implemented in developing countries.

4.4.3 Co-Colonization of Antibiotic Susceptible and Non-Susceptible Strains

An important observation in this study was that 57% (4/7) of multiple colonization episodes had mixed antibiotypes illustrating the potential for genetic exchange between strains with different drug susceptibility. For example, in four colonization episodes there were both penicillin susceptible and non-susceptible strains colonizing together. To our knowledge, the simultaneous carriage of penicillin susceptible together with non-susceptible strains has previously been reported in two other studies (Gratten *et al.* 1989; Sa-Leao *et al.* 2002). Two episodes also had simultaneous colonization of Sxt susceptible and non-susceptible pneumococci, and to our knowledge this has not been previously reported. The spread of penicillin resistance in pneumococci has been attributed to the HGT of *pbp2x*, *pbp2b*, and *pbp1a*, predominantly through mosaic blocks (Dowson *et al.* 1989; Laible, Spratt & Hakenbeck 1991; Hakenbeck *et al.* 1998). Similarly, HGT between pneumococci and viridans streptococci have been proposed to contribute to non-susceptibility to Sxt (Wilén *et al.* 2009). Therefore, a mixed colonization of antibiotic-susceptible and non-susceptible strains increases the potential for susceptible strains to acquire resistant determinants, enabling them to survive the presence of antimicrobials.

4.4.4 Geographically Distinct STs

The semi-closed characteristics of the cohort may also explain the high proportion of geographically distinct STs detected. Nearly 90% of all STs detected have thus far been exclusively found in African strains. However, it is likely that the MLST database is not a representation of the true ST diversity of pneumococci in the sub-Sahara region, as the majority of strain information originates from developed countries. Nonetheless, most of the STs detected, including novel STs, are SLVs of pre-existing strains detected in Africa. Furthermore, some of these STs in Tanzania form a CC of SLVs exclusively detected in the continent (Appendix Table A3). Geographically unique STs were also found in Latin America (Reis *et al.* 2008;

Inverarity *et al.* 2011) and Asia (Jefferies, Tee & Clarke 2011). Similar to our cohort, some of these studies were conducted in areas with high-density populations such as urban slums and refugee camps. Donkor *et al.* (Donkor *et al.* 2011) postulates that high prevalence of multiple carriage directly contributes to emergence of novel ST by the HGT of allelic variants between co-colonizing strains. Additionally, we propose that the extent of diversity within the nasopharynx also influences emergence of novel STs. However, this premise assumes that the MLST database for sub-Sahara is an accurate representation of the population in this region. At present approximately 70% of STs present in the database corresponds to pneumococci in the western world (Donkor *et al.* 2011). Additional genotypic characterization of pneumococci should therefore focus on strains isolated in other geographical regions.

4.4.5 Detection of NT Strain

A NT strain, based on absence of agglutination when mixed with anticapsular antibodies, was detected. This strain may be an acapsulate strain, as *cpsA*-targeting primers developed previously (Pai, Gertz & Beall 2006) provided no amplicons. However, up to 2% of encapsulated pneumococci may not have *cpsA* amplified using this method. This strain is bile-soluble, and the MLST loci in this strain (alleles of ST4156) share 100% homology to pre-existing MLST loci in the database, strongly suggesting that this strain is a pneumococcus. NT and acapsulate strains are commonly isolated from the nasopharynx, and studies have shown that they are genetically-diverse, antimicrobial-nonsusceptible, and may act as potential sources of genetic materials encoding adaptive traits (Hauser, Aebi & Mühlemann 2004; Hanage *et al.* 2006; Sá-Leão *et al.* 2006; Andrade *et al.* 2010; Marsh *et al.* 2010; Simões *et al.* 2011b). The multi-resistant Norwegian clone ST344 is among the most common NT clone in the nasopharynx and infection sites in Europe (Sá-Leão *et al.* 2006; Simões *et al.* 2011b). ST4156 from this study share no identical allele with ST344. According to the MLST database, ST4156 shares 4/7 alleles with numerous encapsulated STs (ST4662, 5109, 5228, 6772) found in China and Thailand. As capsule is a barrier to genetic transformation, acapsulate strains may be more efficient in DNA uptake (Weiser & Kapoor 1999; Pearce, Iannelli & Pozzi 2002). The NT strain (ST4156) detected in our study was co-colonizing with an unrelated serotype 6B strain (ST4429), sharing only 1/7 loci (*aroE*). Based on a previous study, high proportions of NT and acapsulate strains were found co-colonizing with capsulated strains of

different serotypes, indicating that these strains possibly play crucial roles in pneumococcal evolution (Brugger *et al.* 2010).

4.4.5 Novel STs by HGT During Co-Colonization

A high proportion (58%) of STs detected were novel STs, and the study suggests that some of these novel STs may have arisen through HGT during co-colonization; these strains are SLVs and DLVs of pre-existing STs but the locus (or loci) that differed were found to be carried by co-colonizing strains. It has been postulated that there is a greater potential for new STs to arise if the carriage rate of multiple strains is higher (Donkor *et al.* 2011). Although, in this study, the carriage of multiple strains was 12%, lower than the 19% reported by Donkor's group, the diversity of strains within the multiple colonization events of our cohort was greater, which would also increase the potential for generating new STs. While SLVs differing in single nucleotides may have evolved by mutational events (Feil, Enright & Spratt 2000; Feil *et al.* 2004), SLVs containing multiple nucleotide differences within a locus are likely to have arisen through HGT. It is not known how these new clones would contribute to the supragenome pool and pneumococcal evolution over time, as a high proportion of detected STs may be transient clones and disappearing due to competition by more successful clones (Jefferies *et al.* 2010a). However, these new STs may adapt to stabilize and survive by interaction with co-colonizing strains.

Despite the serotype diversity observed in multiple colonizations our sequence typing data did not reveal any evidence of capsule switching at the individual colonization episode level as well as the cohort level, as strains of different serotypes also had non-related STs.

4.4.6 Repeated Colonization of Pneumococcal Strain

Sequence typing revealed that the same strain (ST4432) of a 6B serotype was isolated five months later (January and June) in Child 11, and a new strain of serotype 6B was isolated in Child 3 six months later (ST4368 then ST4373). A similar observation was made in a previous longitudinal report (Sluijter *et al.* 1998). In that study, a single colony was characterized monthly from the same child, and different strains of the same serotypes were detected in the same child at different months, suggesting re-acquisition. From an immunological perspective, our data on re-acquisition or co-

acquisition of a different strain with the same serotype may provide further support that colonization does not provide complete immunity, and that protective immunity to colonization is not solely dependent on capsular polysaccharide (McCool *et al.* 2002). This is borne out by the absence of a correlation between a previous colonization by serotype 6B and the titer of serotype 6B-specific antibodies (Weinberger *et al.* 2008). Thus, additional surface proteins may elicit protective immunity to colonization, and these may be accessory genes that are not present in all strains of a particular serotype, or within a CC of related STs, or even within a particular ST (Donati *et al.* 2010). However, caution must be taken for such interpretations. The absence of capturing pneumococci in the interceding months that indicates re-acquisition is inconclusive as these strains may persist below the level of detection by the method employed in this work. The detection limit of the culture method employed is likely to be lower than those of DNA-based multiplex PCR and microarrays. The latter methods may therefore be used on the original stored URT samples to potentially address and provide greater understanding of the pneumococcal colonization dynamics of this cohort.

4.5 Concluding Remarks

The systematic characterization of up to twenty colonies from the nose and throat has revealed the presence of multiple strains co-colonizing. These colonization events contain pneumococci of different serotypes, antibiotic susceptibility profiles, and genetic backgrounds as determined by MLST. MLST information provided evidence indicating assessing serotype diversity alone is an underestimate of true strain diversity. A nontypable strain was detected co-colonizing with an encapsulated strain. A high proportion of novel STs had been identified, and these novel STs may have arisen through horizontal gene transfer with co-colonizing strains.

CHAPTER FIVE: Analysis of United Kingdom and Tanzania Pherotypes in Carriage and Clinical Strains

5.1 Introduction

5.1.1 Overview of Competence

S. pneumoniae is naturally competent in acquiring exogenous DNA from their immediate environment. This state of competence is induced and regulated in by a two-component, quorum-sensing system known as the competence stimulatory pathway. It regulates a set of more than 180 genes grouped as early, late, delayed, and repressed genes in a coordinated manner based on the time of gene regulation in relation to competence activation (Alloing *et al.* 1998; Peterson *et al.* 2000; Rimini *et al.* 2000; Dagkessamanskaia *et al.* 2004; Peterson *et al.* 2004).

The pathway consists of an ATP-binding cassette (ABC) transporter and accessory proteins ComA and ComB, the competence stimulating peptide (CSP, encoded by *comC*), its cognate histidine kinase receptor (ComD), and a response regulator ComE (Pestova, Håvarstein & Morrison 1996; Claverys & Havarstein 2002) (Fig. 5.1). CSP is thought to be processed and exported by ComAB (Hui & Morrison 1991; Hui, Zhou & Morrison 1995). Extracellular accumulation of CSP reaches a strain-specific threshold CSP concentration (usually between 1-10 ng/μL), during which competence is induced by the binding of CSP to the histidine kinase receptor ComD, leading to the autophosphorylation of ComD at the cytoplasmic histidine residue, followed by the phosphorylation of the response regulator ComE (ComE-P). Most pneumococci of a pherotype can be induced to competence only by the same pherotype, however strains with non-cognate receptors have been identified, and some strains may be induced by different pherotypes at increased concentrations (Pozzi *et al.* 1996; Whatmore, Barcus & Dowson 1999; Iannelli, Oggioni & Pozzi 2005).

ComE-P recognizes a region of a promoter containing two 9-bp imperfect repeats separated by 12 bps (Ween, Gaustad & Håvarstein 1999). Sets of early and late genes containing this motif are activated by ComE-P including the *comAB* and *comCDE* genes themselves. This creates a positive feedback loop in the competence activation circuit. ComE-P also upregulates expression of *comX*, which encodes an alternative

sigma factor (Lee & Morrison 1999). ComX associates with the RNA polymerase holoenzyme and affects transcription of genes containing promoters with a *cin*-box consensus sequence (TAATACGAATA) (Campbell, Choi & Masure 1998; Lee & Morrison 1999; Luo, Li & Morrison 2003; Luo & Morrison 2003). ComX regulates the transcription of late genes including ones involved in transformation and recombination as well as stress responses, and expression of these late genes increases approximately 10 minutes after CSP exposure *in vitro* (Dagkessamanskaia *et al.* 2004). The transient nature of the competence state may be mediated by the differential transcriptional responses of ComE depending on its phosphorylation state, possibly mediated by a phosphatase targeting ComE-P (Alloing *et al.* 1998). The CiaRH TCS may also regulate the competence property, as inactivation of *ciaR* was associated with an increase in *comC* expression (Mascher *et al.* 2003).

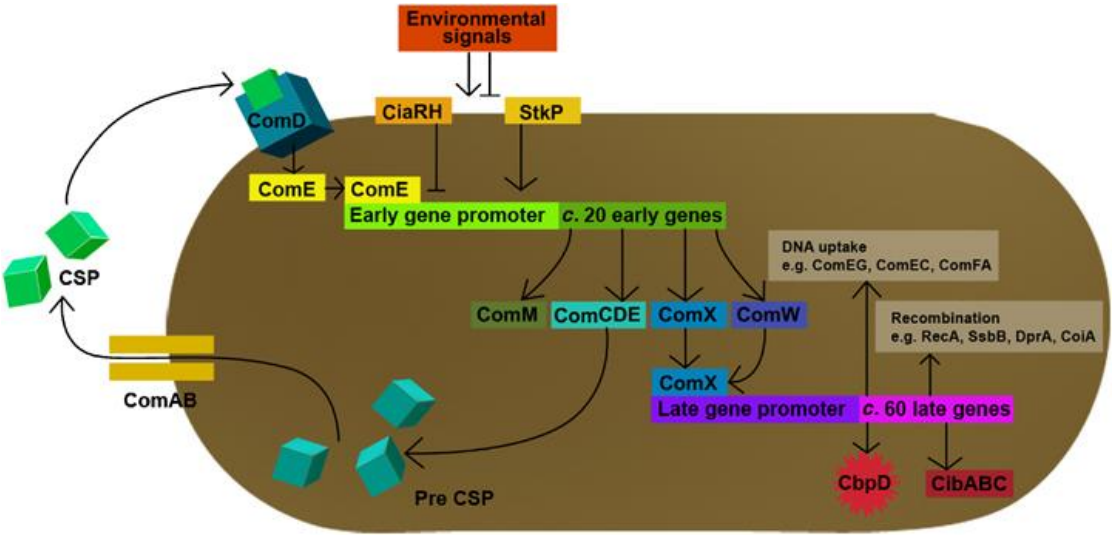


Fig. 5.1. The competence stimulatory pathway in the pneumococcus. CSP is synthesized by *comC* as a pre-peptide containing a 24-aa leader peptide containing a double-glycine motif. The processing, leader cleavage, and export of pre-CSP are mediated by the ABC transporter ComAB, releasing the 17-aa mature peptide into the extracellular space. As CSP is exported and accumulated, a threshold concentration is reached where competence induction is trigger by the binding of CSP to its cognate ComD histidine kinase sensor receptor, leading to autophosphorylation of ComD, in turn phosphorylating the response regulator ComE, leading to the expression of early genes including fratricide immunity protein gene *comM*, alternative sigma factor gene *comX*, and *comW* which encodes a protein responsible for the stability and function of ComX. When ComX associates with RNA polymerase, transcription of late genes with *cin*-box promoters takes place. Activation of late competence genes affects pneumococcal DNA uptake and recombination, biofilm formation, and bacterial fratricide. Figure from Johnsborg and Håvarstein. (Johnsborg & Håvarstein 2009) with permission granted from FEMS Microbiology Reviews via RightsLink Copyright Clearance Center.

5.1.2 Role of Competence in Pneumococcal Physiology

Despite the large number of genes regulated by competence, only 23 are required for transformation, suggesting that the competence pathway genes are involved in additional aspects of pneumococcal physiology (Peterson *et al* 2004). Indeed, competence induction may be associated with virulence, as pneumococci with mutations in early competence genes are attenuated in murine infections in the lungs and blood (Bartilson *et al.* 2001; Lau *et al.* 2001; Oggioni *et al.* 2006). Expression studies have also linked competence activation to growth arrest and the induction of stress response (Dagkessamanskaia *et al.* 2004). In a murine sepsis model, it was shown that the treatment with CSP inhibits pneumococcal growth, and increases survival rate of mice (Oggioni *et al.* 2004). Such observations appeared to be dependent on the activation of ComD.

Recently it was discovered that competence induction actively triggers cell lysis and DNA release in a subpopulation of non-competent cells, mediated by murein hydrolases LytA, LytC (provided by either competent cells or noncompetent target cells) and CbpD, and bacteriocin-like peptides (CibAB) (Steinmoen, Knutsen & Håvarstein 2002; Steinmoen, Teigen & Håvarstein 2003; Guiral *et al.* 2005; Kausmally *et al.* 2005; Eldholm *et al.* 2009; Eldholm *et al.* 2010). Competent cells express an immunity factor, encoded by an early gene, *comM*, to prevent lysis from its own lytic enzymes (Håvarstein *et al.* 2006a). This process, known as fratricide, provides competent cells with sources of external DNA for acquisition, as fratricide has been shown to directly enhance DNA uptake (Johnsborg *et al.* 2008). Acquisition of these DNA may provide novel traits with adaptive benefits. Also, fratricide has been proposed to act as virulence complementation between co-colonizing strains by releasing virulence factors from heterogeneous pneumococci, thereby modulating host immune response (Guiral *et al.* 2005). It is interesting to note that interspecies fratricide also occurs, in agreement to the hypothesis that the pneumococcal supragenome (Section 1.6.3) may include genes of *S. mitis* and *S. oralis*. Activation of competence may confer a fitness cost, as *comE* mutants colonized better than wild type strains in rat colonization competition assays (Kowalko & Sebert 2008).

5.1.3 Role of Competence in Colonization and Biofilm Formation

DNA release from competence-mediated fratricide may support the biofilm structure, as exogenous DNA is a constituent of biofilm, and treatment of DNase is associated with reduced biofilm biomass and impaired biofilm formation (Moscoso, García & López 2006; Hall-Stoodley *et al.* 2008). While no direct mechanism has been described for the precise role competence plays in biofilm, cells adopting a biofilm mode of growth have increased expression of *comA, comE,* and *comX* (Oggioni *et al.* 2006). As transformation, fratricide, and biofilm formation are regulated by competence, this transient physiological state may provide the appropriate conditions for genetic exchange and DNA acquisition during pneumococcal colonization. A comparison of three different *in vitro* models of biofilm growth revealed different roles of competence pherotype in biofilm formation: while competence is crucial for biofilm maintenance in biofilms studied in microtiter plates, its role in biofilms grown in continuous culture is minimal (Trappetti *et al.* 2011a).

Formation of a biofilm extracellular matrix was recently shown to require the expression of LuxS, encoding S-ribosylhomocysteine lyase, and the presence of Fe (III). This is concomitant with the upregulation of iron uptake transporter PiuA, the early and late competence genes, increase in transformation efficiency, release of extracellular DNA, and fratricidal lysis of noncompetent cells (Trappetti *et al.* 2011c).

.

5.1.4 The Role of Competence in Biofilm Formation in Other Streptococci

The link between competence and biofilm formation in streptococci was first identified in *S. gordonii,* where inactivation of *comD* by transposon mutagenesis rendered mutants defective in biofilm formation (Loo, Corliss & Ganeshkumar 2000). Similarly, competence is associated with biofilm formation in *S. mutans*, a predominant causative agent of dental caries. In this sense, competence in *S. mutans* is related to its pathogenicity, because biofilm formation allows for persistence of the organism from rapidly changing environmental conditions and physical stress of the oral cavity (Li *et al.* 2001). It is postulated that competence also affects the initial stages of biofilm formation in *S. mutans,* as *comD* mutants affected *in vitro* adherence to mucin-coated plate surfaces in this organism (Li *et al.* 2002). In a mechanism possibly similar to pneumococcal fratricide, competence induction is associated with cell death, chromosomal DNA release and increase in biofilm biomass in *S. mutans*

(Perry, Cvitkovitch & Lévesque 2009) The addition of synthetic CSP was also demonstrated to enhance biofilm formation in *S. intermedius* (Petersen, Pecharki & Scheie 2004).

5.1.5 The Competence Stimulatory Peptide – ComC

As early as the 1960s it was known that competent pneumococci secrete a competence factor, of a proteinaceous nature, as exposure of this factor in trypsin abolishes competence induction (Tomasz & Hotchkiss 1964). In 1995, the chemical nature of this factor was described by Håvarstein *et al.* (Håvarstein, Coomaraswamy & Morrison 1995). By chromatographic isolation of competent culture supernatants, the competence stimulatory peptide (ComC) was purified, and its amino acid sequence determined. Nucleotide sequence analysis revealed that ComC is synthesized as a 41-amino acid pro-peptide containing a 24-aa leader sequence ending with a Gly-Gly motif. The leader sequence also contains the bacteriocin processing consensus residues with hydrophobic residues at positions -4, -7, -12, and -15. The homology of the leader peptide to other double-glycine-containing bacteriocins (Håvarstein, Holo & Nes 1994; Dirix *et al.* 2004), and the identification that ABC transporters process and export bacteriocins with double-glycine motifs (Håvarstein, Diep & Nes 1995), led to the designation of ComA and ComB as the ABC transporter complex responsible for cleavage of leader peptide of CSP and the export of the 17-aa mature peptide (Håvarstein, Diep & Nes 1995; Hui, Zhou & Morrison 1995).

5.1.6 Diversity of ComC in the Pneumococcus

Two major CSP pherotypes, CSP1 and CSP2, were identified (Pozzi *et al.* 1996). Sequence comparison shows that these two alleles differ by eight amino acid changes in the mature peptide region. Later studies identified additional rare CSP alleles, with up to 7 identified in the natural population (Fig. 5.2) (Ramirez, Morrison & Tomasz 1997; Whatmore, Barcus & Dowson 1999; Vestrheim *et al.* 2011). Differences in amino acids between these pherotypes include residues that confer biological activity (Havarstein 1999; Johnsborg *et al.* 2006). Also, substitution of residues 4, 7, 8, and 12 (together form a hydrophobic patch in the ligand) of mature CSP2 peptide to the corresponding amino acids in CSP1 shifts the ligand specificity to the receptor ComD1, suggesting that these variable amino acid residues mediate receptor specificity. Studies investigating the diversities of CSP are listed in Table 5.1.

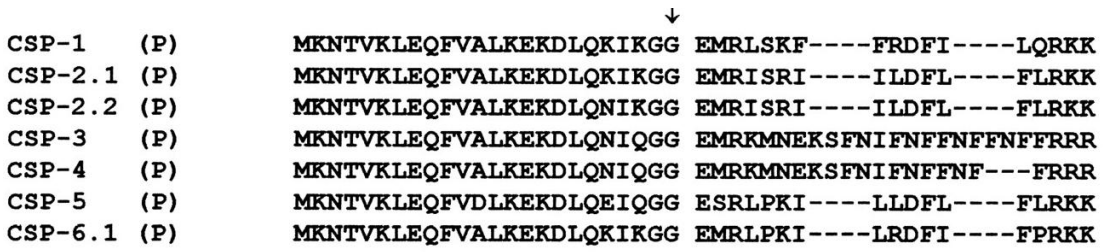


Fig. 5.2 Amino acid alignment of pneumococcal pherotypes. Arrow indicates position of double-glycine residues cleaved during CSP export. CSP6.1 was recently hypothesized to be a common allele of *S. pseudopneumoniae* (Chapter 7). Figure from Whatmore *et al.* (Whatmore, Barcus & Dowson 1999).

Table 5.1. Selected publications describing pneumococcal pherotype diversity

Reference	No. Strains	Geographical Location	Colonization or Infection	Pherotype	Proportion (%)
(Pozzi *et al.* 1996)	42	Italy, USA	Infection	CSP1 CSP2	29 (69.0) 13 (31.0)
(Ramirez, Morrison & Tomasz 1997)	50	USA South Africa South Korea Spain Croatia Iceland Hungary Czech Republic	Both	CSP1 CSP2.1 CSP2.2 CSP3	24 (48.0) 16 (32.0) 9 (18.0) 1 (2.0)
(Whatmore, Barcus & Dowson 1999)	60	USA Spain Papua New Guinea UK Kenya Czechoslovakia Mexico France	Unknown	CSP1 CSP2.1 CSP2.2 CSP3 CSP4 CSP5 CSP6.1	29 (48.3) 20 (33.3) 7 (11.7) 1 (1.7) 1 (1.7) 1 (1.7) 1 (1.7)
(Sa-Leao *et al.* 2002)	8[a]	Portugal	Colonization	CSP1 CSP2.1 CSP2.2	5 (62.5) 2 (25.0) 1 (12.5)
(Carrolo *et al.* 2009)	483	Portugal	Infection	CSP1 CSP2	341 (70.6) 142 (29.4)
(Cornejo, McGee & Rozen 2010)	88	Various[b]	Infection	CSP1 CSP2	65 (73.9) 23 (26.1)
(Vestrheim *et al.* 2011)	150	Norway	Colonization	CSP1 CSP2	94 (62.7) 56 (37.3)

[a] Analyzed co-colonization of multiple pherotypes, but small sample size

[b] The 88 strains are a sample from a larger collection of strains from South Africa, USA, Sweden, Spain, New Zealand, Taiwan, Argentina, Brazil, Hong Kong, and France (Yu *et al.* 2003).

ComC is also present in other streptococci (Håvarstein, Hakenbeck & Gaustad 1997; Allan *et al.* 2007; Kilian *et al.* 2008). Switching of competence pherotypes has been reported to take place between streptococcal species (Håvarstein, Hakenbeck & Gaustad 1997; Kilian *et al.* 2008). The recombination breakpoint extends beyond regions of *comC* and *comD* mediating pherotype specificity, indicating that strains can change both their pherotypes and receptor by this manner without resulting in non-cognate ligand-receptor pairs. Extensive diversity of ComC between streptococcal species has been observed in a recent study (Kilian *et al.* 2008), where each strain of *S. mitis* and *S. pseudopneumoniae* analyzed possessed a distinct pherotype. Chapter 7 describes the pherotype diversity and sequences associated with *S. pseudopneumoniae*. The presence of multiple pherotypes is also detected in *S. mutans* (Allan *et al.* 2007). *In vitro* cross-induction of competence was seen with two pherotypes of *S. sanguinis* containing multiple amino acid substitutions in the mature CSP peptide (Zhu *et al.* 2011). The diversity of CSP therefore does not appear to be a property exclusive to *S. pneumoniae*, but a common feature of the Streptococcus genus.

5.1.7 Significance of Pherotype Diversity

The full biological role of pherotype diversity remains to be elucidated. One hypothesis argues that pherotypes drive genetic differentiation between strains within a species; the release of CSP of one pherotype in most pneumococci will only induce competence and hence genetic transformation of the matching pherotype. Strains belonging to a given pherotype are significantly associated with other properties such as serotypes and antibiotic resistance, and even more so with genetic backgrounds as measured by PFGE and MLST (Carrolo *et al.* 2009; Vestrheim *et al.* 2011). Håvarstein (Håvarstein, Hakenbeck & Gaustad 1997) argues that pherotype specificity allows for communication within bacteria of the same pherotype rather than between different pherotypes. Indeed, with pherotype-specific competence induction, only a subset of cells of the same pherotype would undergo competence induction and DNA acquisition at one time in general. As DNA uptake in the *S. pneumoniae* does not require a signal sequence, pherotype specificity may limit acquisition of highly divergent sequences, which may be deleterious or have toxic effects (Tortosa & Dubnau 1999). Therefore, while acknowledging the phenomenon

of inter-pherotype genetic exchange, Carrolo *et al.* (Carrolo *et al.* 2009) hypothesizes that this is relatively rare compared to intra-pherotype exchanges.

A second hypothesis suggests that recombinational events are not limited to strains bearing identical pherotypes, because bacterial fratricide would enable the release of DNA from non-competent cells (Cornejo, McGee & Rozen 2010). If genetic differentiation occurs as a result of pherotype differences, sequence diversity of strains within a pherotype should be less across pherotypes. By comparing MLST sequences of CSP1 and CSP2, Cornejo *et al*. (Cornejo, McGee & Rozen 2010) did not find this to be the case. Pneumococcal competence can lead to the lysis of non-competent pneumococci as well as species of viridans streptococci (Johnsborg *et al.* 2008). Therefore, by inducing competence only in a subset of cells of a pherotype, gene exchanges between pherotypes may actually be promoted by allowing the lysis and DNA uptake of non-competent cells of different pherotypes.

5.1.8 Cross-Activation of Competence by Non-Cognate Pherotypes

While most competent pneumococci could only be activated by their own CSP allele, the work of Pozzi *et al.* (Pozzi *et al.* 1996) demonstrated that some strains can be activated by both CSP1 and CSP2 in experimental conditions. Cross-activation work was also performed on strains expressing comD3 and comD4 in which some of these strains were observed to also be activated by the CSP1 ligand at higher CSP concentrations (300-1000 ng/mL) (Iannelli, Oggioni & Pozzi 2005). To date, cross-activation studies have been performed *in vitro*, but no *in vivo* results have been reported. Similarly, it is not known whether at physiological conditions, the concentration of CSP ligands reach a level that enable cross-activation. Should cross-activation occur in natural populations, understanding the pherotype diversity of co-colonizing pherotypes would provide additional insight as to the communication between and dynamics of co-colonizing strains. To our knowledge, only two studies had been conducted to investigate pherotype diversity in co-colonizations (Sa-Leao *et al.* 2002; Vestrheim *et al.* 2011).

5.1.9 Aim

Multiple pneumococcal pherotypes exist, however there is scant data on the virulence of different phenotypes, and the diversity of pherotypes had only been reported in a small number of studies. Similarly, information regarding the relationship between pherotypes and core genetic backgrounds, and the co-colonization of different pherotypes is limited. In this study, the pneumococcal pherotypes from carriage (Chapter 4) and invasive strains obtained from Tanzanian children and patients respectively were compared with UK strains. Comparison of pherotype prevalences between carriage and clinical pneumococci from the UK and Tanzania will allow us to determine if differences in proportions of pherotypes are observed between these groups, and whether these pherotypes differ in disease potential and geographical distribution. Examining pherotype diversity of pneumococci in co-colonization events as described in Chapter 4 provides insights as to how these events provide a suitable condition for genetic exchanges.

5.2 Materials and Methods

5.2.1 Bacterial Strains

CSP analysis was performed on a total of 200 pneumococcal strains. Of these 115 were isolated between the years 1993-2000 from the Royal Free Hospital in London, U.K. These included 68 from blood, 14 from the lower respiratory tract, 4 from the middle ear, 2 from cerebrospinal fluid, and one from conjunctiva. Twenty-six strains were isolated from unknown anatomical sites due to loss of clinical data. Eighty-five pneumococci from Tanzania, including 18 invasive strains (Crump *et al.* 2011a; Crump *et al.* 2011b), and 67 carriage strains from healthy children (Chapter 4). Please refer to Appendix Table A1 for strain included in this study.

5.2.2 PCR of *comC*

Sequences for primers used (*comC-fw* and *comC-rv*), and the annealing temperature and extension time of the PCR are indicated in Table 2.6.

5.2.3 Pherotype Characterization by RFLP

For 39 of the 85 strains from Tanzania, *comC* amplicon (~300 bp) was subjected to restriction analysis by endonuclease *Bsa*XI (New England Biolabs). This was performed by Dr. Sarah Batt prior to the commencement of works included in this thesis. Upon restriction, the largest band for CSP1, 2, and 4 were expected to be ~150 bp, ~200 bp, and 300 bp (undigested), respectively, based on *in silico* restriction on NEBcutter (Vincze 2003) using published *comC* sequences as template (accession numbers AJ240789, AJ240785, AJ240765, and AJ240792. Strains typed as CSP2 on RFLP was subtyped as CSP2.1 and 2.2 by sequencing. Sequencing was also performed on strains with undigested band to confirm pherotype. All strains tested as undigested by RFLP were CSP4 by sequencing (see below).

5.2.4 Pherotype Characterization by *comC* Sequencing

Gel electrophoresis, DNA purification, cycle sequencing, ethanol precipitation, and sequence analysis were performed as outlined in Sections 2.8.3 to 2.8.8.

5.2.5 Statistical Analysis

Statistical analyses were performed on Graph Prism 5 (GraphPad) and SPSS Version 20. Fisher's Exact Test was employed to compare significances of differences in

pherotype proportions. Odds ratio for likelihood of pherotypes to be associated with geographical region and invasiveness was as calculated. The odd of detecting each pherotype in Tanzania relative to the United Kingdom was calculated by: *ratio = (ad)/(bc)*, where *a* is the number of Tanzania strains of the pherotype in question, *d* is the number of United Kingdom of other pherotypes, *b* is the number of United Kingdom strains of the pherotype, and *c* is the number of Tanzania strains of other pherotypes. For disease potential, the odds ratio of disease relative to carriage is calculated using the same formula, where *a* is the number of invasive strains of the pherotype in question, *d* is the number of carriage strains of other pherotypes, *b* is the number of carriage strains of the pherotype, and *c* is the number of invasive strains of other pherotypes. An odds ratio of >1 indicates an increased likelihood for a pherotype to be associated with Tanzania (for geography) or disease (disease potential), where odds ratio of <1 indicates a reduced likelihood.

5.3 Results

5.3.1 Pneumococcal pherotypes from the UK and Tanzania

Pneumococcal pherotypes were determined for 200 strains. Of these, a hundred and fifteen were United Kingdom clinical pneumococcal strains, eighteen were Tanzania invasive strains, and 67 were Tanzanian nasopharyngeal strains from healthy children under 6 years of age (Chapter 4). Four pherotypes were identified, CSP1 77.4% (89/115), 13.0% were CSP2.1 (15/115), 5.2% were CSP4 (6/115), and 4.4% were CSP2.2 (5/115).

The prevalence of CSP1 (41.2%, 35/85) was significantly lower in Tanzanian strains compared with U. K. strains (Table 5.2). Conversely, CSP2.1 (30.6%, 26/85) was also significantly more prevalent in Tanzania. CSP2.2 was not detected in Tanzania (Table 1). The CSP4 pherotype Tanzania (27.1%, 23/85) was greater by more than five-fold. A single CSP3 pherotype was also detected in a Tanzanian strain isolated from blood.

Table 5.2. Odds ratio for increased likelihood of pherotypes association by geography and disease potential.

Variable	Pherotype (%)				
	CSP1	**CSP2.1**	**CSP2.2**	**CSP3[b]**	**CSP4**
Geography					
Tanzania (n=85)	35 (41.2)	26 (30.6)	0 (0)	1 (1.2)	23 (27.1)
United Kingdom (n=115)	89 (77.4)	15 (13.0)	5 (4.4)	0 (0)	6 (5.2)
Tanzania:UK ratio	0.20	2.94	0.12	4.10	6.74
95% CI	(0.11-0.38)	(1.44-5.99)	(0.01-2.16)	(0.16-102.0)	(2.60-17.45)
P-value[a]	**< 0.0001**	**0.0042**	0.074	0.43	**<0.0001**
Disease Potential					
Tanzania Disease (n=18)	9 (50.0)	5 (27.8)	0 (0)	1 (5.6)	3 (16.7)
Tanzania Carriage (n=67)	26 (38.8)	21 (31.3)	0 (0)	0 (0)	20 (29.9)
(Disease:Carriage ratio)	1.58	0.84	Not detected	11.57	0.47
(95% CI)	(0.55-2.24)	(0.27-2.67)	in Tanzania	(0.45-296.7)	(0.12-1.81)
P-value[a]	0.43	1.00	-	0.21	0.37

[a] P-value calculated by two-sided Fisher's Exact Test. Statistical significance attained when P-value < 0.05. Statistically-significant comparisons in bold

[b] CSP3 not detected in the United Kingdom

5.3.2 Pherotype and Antibiotic Susceptibility in Tanzanian Strains

Analysis of antibiotic susceptibility data for penicillin and co-trimoxazole data that was available for 44 Tanzania strains revealed no association between pherotype and drug susceptibility (Table 5.3).

Table 5.3. Antibiotic susceptibilities of penicillin and co-trimoxazole in Tanzanian strains by pherotype

	CSP1 (% of n=18)		CSP2.1 (% of n=12)		CSP3 (% of n=1)		CSP4 (% of n=13)		Total (% of n=44)	
Antibiotic[a]	Susceptible	Non-susceptible	Susceptible	Non-susceptible	Susceptible	Non-susceptible	Susceptible	Non-susceptible	Susceptible	Non-susceptible
Penicillin	12 (67)	6 (33)	3 (25)	9 (75)	1 (100)	0 (0)	4 (31)	9 (69)	20 (45)	24 (55)
Co-trimoxazole	6 (33)	12 (67)	2 (17)	10 (13)	0 (0)	1 (100)	6 (46)	7 (54)	14 (32)	30 (68)

[a] Penicillin nonsusceptibility determined as minimum inhibitory concentrations ≤0.06 μg/mL using E-test; co-trimoxazole nonsusceptibility determined as a zone of inhibition ≥ 19 mm by the Kirby-Bauer disc diffusion method.

5.3.3 Invasive disease potential for individual pherotypes

To ascertain if particular pherotypes are associated with invasive disease, the invasive odds ratio for each pherotype was determined on 67 carriage strains and 18 invasive strains from Tanzania. No significant association was found (Table 5.2). The prevalence of the CSP1 pherotype in U.K. clinical strains is significantly greater than that of clinical strains in Tanzania, while significance within clinical strains from the two countries was not observed for the other pherotypes (Table 5.4).

Table 5.4. Odds ratio for increased likelihood of clinical pherotypes association by geography

Country	**Pherotype (%)**				
	CSP1	**CSP2.1**	**CSP2.2**	**CSP3**	**CSP4**
Tanzania (n=18)	9 (50.0)	5 (27.8)	0 (0)	1 (5.6)	3 (16.7)
UK (n=115)	89 (77.4)	15 (13.0)	5 (4.4)	0 (0)	6 (5.2)
Tanzania:UK Ratio	0.29	2.13	0.54	19.80	3.63
95% CI	(0.11-0.81)	(0.88-5.15)	(0.029-10.24)	(0.77-505.9)	(0.82-16.08)
p-value	**0.021**	0.15	1.00	0.14	0.10

5.3.4 Pherotypes of Co-Colonizing Strains

Previously we identified seven colonization events with multiple pneumococcal strains as characterized by serotype and sequence type (ST) by MLST (Chapter 4). Pherotype analysis of these co-colonizing strains revealed that all but one event had multiple pherotypes (Table 5.5). Most co-colonization events had two pherotypes, while three pherotypes were detected in colonizations 2 and 3 (Table 5.5). Within a host, strains with the same serotype, but different STs also had different pherotypes (e.g. serotype 6B strains associated with CSP2.1 and 4. Strains with different serotypes and/or STs were found to be of the same pherotype (e.g. colonizations 3 and 5, Table 5.5).

Table 5.5. The serotypes, pherotypes, sequence types, and MICs of penicillin and co-trimoxazole present in Tanzanian colonizing strains

Sample	Child (Month)[a]	Serotype[b]	CSP Pherotype	Sequence Type (ST)
1	1	NT	4	4156
		6B	2.1	4429
2	16	19F	1	347
		34	2.1	4158
		35A	4	840
3	35	6B	2.1	854
		6B	4	4157
		10A	1	852
		21	1	1145
		19A	2.1	4162
		19F	1	6170
4	3 (Sep)	1	1	217
		6B	4	4373
5	29	35A	4	840
		11A	4	5752
6	10 (Jan)	19A	2.1	4162
		13	4	4370
7	10 (Apr)	18C	1	1233
		6B	4	4368
8	18	10A	1	852
9	22	17F	2.1	4160
10	45	21	1	1145
11	3 (Jan)	6B	4	4368
12	3 (Mar)	6B	4	4368
13	24	19A	2.1	4162
14	3 (Jul)	1	1	217

[a] The months of colonization event are indicated if a child had multiple co-colonization events during study period

[b] NT: nontypable by agglutination and absence of amplification targeting *cpsB* and *cpsA* (Pai, Gertz & Beall 2006; Leung *et al.* 2011)

5.3.5 Pherotypes of Genetically Related Strains

An eBURST representation of the co-colonization cohort revealed 18 STs (Fig. 5.3). Six of these belong to two pairs of SLVs (ST347 and ST6170 differing at *spi* by one nucleotide, ST4157 and ST4368 differing at *recP* by three nucleotides) and a pair of DLV (ST854 and ST4429 differing at *gki* and *ddl* at five and 40 nucleotides respectively). Each pair of related STs were associated with the same serotype and pherotype (Table 5.5 and Fig. 5.3).

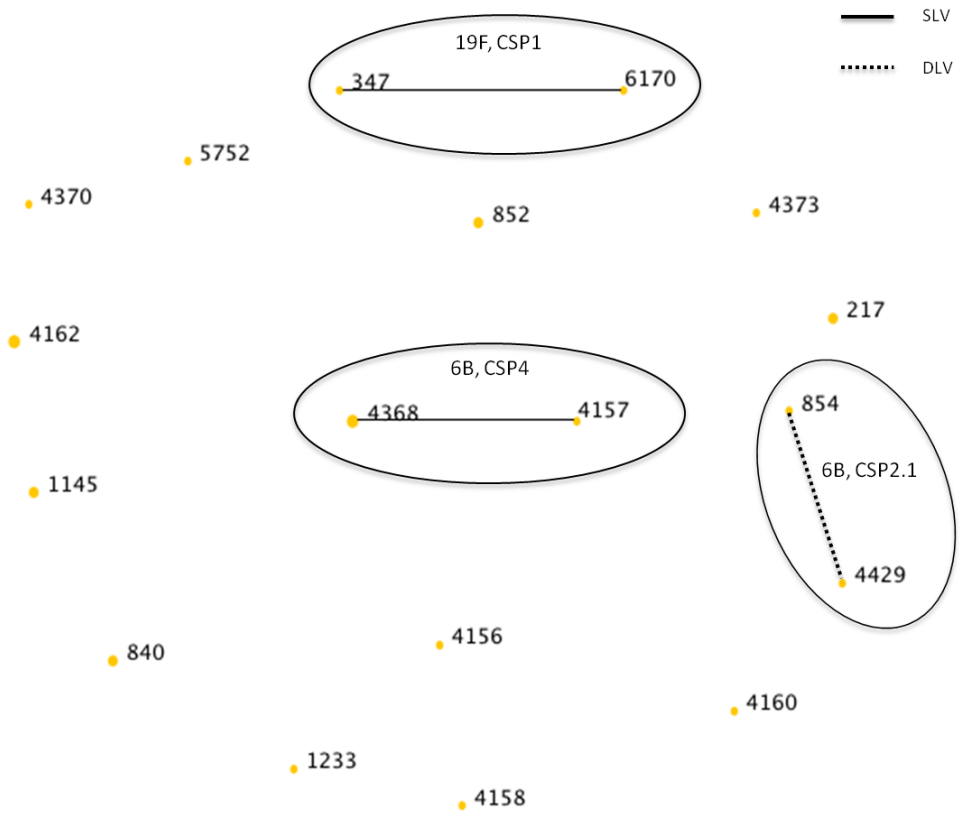


Fig. 5.3. eBURST representation of genetic relatedness of STs and pherotypes of related STs detected in multiple colonization. Genetically related STs with variation in a single allele (single-locus variants) are connected by solid lines, and those differing in two alleles (double-locus variants) are connected by dotted lines. Each pair of related STs encircled and the sharing serotypes and pherotypes indicated.

5.4 Discussion

It is hypothesized that pherotype specificity enables communication between cells within a pherotype, and that this may result in genetically distinct clusters of pneumococci within pherotypes (Håvarstein, Hakenbeck & Gaustad 1997; Carrolo *et al.* 2009). Cellular communication via the activation of competence stimulatory has the potential to propagate genomic plasticity by coordinating genetic transformation (Peterson *et al.* 2004), contributing to biofilm formation (Trappetti *et al.* 2011a; Trappetti *et al.* 2011c), and enabling the release of DNA from non-competent cells through bacterial fratricide (Guiral *et al.* 2005; Johnsborg *et al.* 2008; Eldholm *et al.* 2009). In this study we have investigated the relationship between pherotypes and geography, invasive potential. We also determined the diversity of CSP alleles in co-colonization events.

5.4.1 Prevalence of CSP4

This is the first population-based investigation of pherotype distributions between two different geographical locations. To date, pherotype CSP4 had only been detected in one pneumococcal strain from Papua New Guinea (Whatmore, Barcus & Dowson 1999). However, we found that the CSP4 pherotype was detected in 5.2% of U.K. strains and the prevalence of this pherotype was significantly higher in Tanzania (27%). Interestingly, this pherotype was detected in both carriage (30%) and invasive (17%) Tanzanian strains. Moreover, CSP4 was found to be associated with multiple serotypes and clones in Tanzania, indicating that this pherotype is widely dispersed in this region.

The reasons for the greater prevalence of CSP4 in Tanzanian strains compared with U.K. strains are unknown. One possibility is that there is limited transmission of these strains, as carriage strains were collected from Tanzanian children living in a semi-closed community on a sugar plantation (Chapter 4). Previous studies by us and others have shown that some STs are geographically unique (Reis *et al.* 2008), (Donkor *et al.* 2011; Inverarity *et al.* 2011; Leung *et al.* 2011). The Tanzanian strains with the CSP4 pherotype were associated with STs that have not been detected elsewhere thus far. Conversely, there was a significantly greater prevalence of the CSP1 pherotype in U.K. strains compared with Tanzanian strains. These data indicate that differences in the prevalence of pherotypes between the U.K. and Tanzania may

be due to geographically unique sequence types. Additional characterization of pherotype distribution in sub-Sahara is therefore encouraged.

5.4.2 Association Between Pherotypes and Serotypes, STs, and CCs

Pneumococci of the same serotype were associated with different pherotypes, and in agreement with other studies, genetically related strains as indicated by the same or related STs were found to have the same pherotype (Carrolo *et al.* 2009; Vestrheim *et al.* 2011). Multiple alleles of another virulence factor, pneumolysin, also appear to be clonally distributed, with some alleles postulated to account for strain-specific differences in disease potentials (Kirkham *et al.* 2006; Jefferies *et al.* 2007b; Jefferies *et al.* 2010b; Harvey *et al.* 2011). *In vitro* and *in vivo* models employing isogenic mutants of different CSP variants may increase our understanding of pherotype-specific invasiveness.

5.4.3 Co-Colonization of Different Pherotypes

While simultaneous colonization by different pneumococcal strains has been well documented (Meats *et al.* 2003; Vestrheim *et al.* 2008b; Donkor *et al.* 2011; Leung *et al.* 2011), there are limited studies on co-colonization by different pneumococcal pherotypes (Sa-Leao *et al.* 2002; Vestrheim *et al.* 2011). We have identified that six of the seven (86%) of multiple colonization events contained mixed pherotypes. This is higher than the previous two studies reporting 25% (Sa-Leao *et al.* 2002) and 47% (Vestrheim *et al.* 2011) of multiple colonizations containing mixed pherotypes. We also detected with up to three pherotypes colonizing simultaneously. Under the assumption that multiple colonization allows for genetic exchanges between strains, it appears that the high prevalence of multiple pherotype colonization supports the fratricide hypothesis, where a subpopulation of strains within a colonization event could induce competence and acquire DNA from co-colonizing heterogeneous pherotypes by cell lysis. However, additional large-scale studies are required to further support or disprove the two hypotheses regarding the significance of pherotype diversity (Section 5.1.7). It is possible that both hypotheses contribute to the evolution and diversification of not only pneumococci but other members of streptococci (Section 7.4.5).

5.4.4 Pherotype Cross-Talk

A single CSP3 strain was detected in an invasive strain from Tanzania. Pneumococci with *comC3* were detected in South Africa (Ramirez, Morrison & Tomasz 1997) and Spain (Whatmore, Barcus & Dowson 1999). Interestingly, Iannelli *et al.* (Iannelli, Oggioni & Pozzi 2005) detected *comC1* strains in Italy linked to either *comD3* or *comD4* gene, and that pneumococci with *comD3 or comD4* could respond to increased concentrations (100 ng/mL) of CSP1 and CSP2 *in vitro*. However, whether such CSP concentrations and subsequent cross talks between pherotypes occur in physiological conditions are unknown.

It must be noted that factors other than pherotype, such as capsulation, serotype, ions, antimicrobials, pH, presence of sugars, and other pneumococcal proteins may also affect competence activation and genetic transformability (Yother, McDaniel & Briles 1986; Hsieh *et al.* 2006; Prudhomme *et al.* 2006) reviewed in (Havarstein 1999; Claverys & Havarstein 2002; Johnsborg & Håvarstein 2009). In addition, CSP can be detected in other oral streptococci, including the newly recognized emerging pathogen *Streptococcus pseudopneumoniae* (Kilian *et al.* 2008; Leung *et al.* 2012), and that competence-mediated fratricide in pneumococci can trigger DNA release from related streptococcal species. Therefore, the effect of multiple pherotype colonization on genetic exchange and the adaptation and evolution of *S. pneumoniae* is likely to be as a result of complex interactions between co-colonizing pneumococci, oral streptococci, and the host environment.

5.5 Concluding Remarks

In conclusion, significant geographical differences in the distribution of pherotypes were observed. We reported a higher prevalence of CSP4, contrary to previously reports. Genetically related strains had the same pherotype as well as the same serotype. There was no association found between the pherotype and invasive potential. This study revealed that the majority of co-colonizing strains had different pherotypes, thus providing an optimum natural condition for genetic exchanges between different strains.

CHAPTER SIX: Intra-Serotype Diversity of Pneumococcal Capsular Regulatory Gene *cpsB*

6.1 Introduction

6.1.1 Capsule Regulatory Genes

The first four genes of the *cps* locus (*cpsA, cpsB, cpsC, cpsD*) are expressed in all serotypes except for 3 and 37. CpsA has been shown to bind to the *cps* promoter and deletion of *cpsA* reduces capsule production, and higher levels of *cpsA* expression were detected in serotypes associated with invasive diseases (Morona *et al.* 2000; Bender, Cartee & Yother 2003; Hathaway, Bättig & Mühlemann 2007; Hanson, Lowe & Neely 2011). CpsD is a cytoplasmic protein-tyrosine kinase with autophosphorylating activity, containing Walker A and B ATP-binding domains, and a tyrosine-rich C-terminal with two to four [YGX] repeats, both of which appear to be crucial for capsule synthesis (Morona *et al.* 2000; Morona *et al.* 2003; Morona *et al.* 2004). CpsC is proposed to be a membrane protein required for the phosphorylation of CpsD and attachment of capsule polysaccharides to the cell wall (Morona *et al.* 2000; Bender & Yother 2001; Morona, Morona & Paton 2006; Byrne *et al.* 2011).

6.1.2 CpsB Phosphotyrosine Phosphatase

CpsB is a manganese-dependent phosphotyrosine phosphatase that dephosphorylates CpsD-P and inhibits transphosphorylation between CpsD molecules, thereby activating capsule expression (Morona *et al.* 2000; Morona *et al.* 2002; Morona, Morona & Paton 2006). While deletion of *cpsB* resulted in high levels of CpsD-P in both strains, Rx1-19F had a rough appearance with decreased CPS, while D39 had a smooth appearance with increased CPS (Bender, Cartee & Yother 2003), suggesting that possible additional strain-specific factors located outside the *cps* regulatory region may influence the effects of CpsD phosphorylation on CPS production, and the role of CpsB in modulating capsule production. Recently, CpsC and CpsD has been demonstrated to be crucial for the attachment of capsule material to the cell wall; while deletion of *cpsB*, and hence a build-up of CpsD-P, is associated with decrease in total CPS, approximately 1% of these mutants show a mucoid phenotype due to additional amino acid substitutions, possibly at the extracellular domain of CpsC,

associated with decreased ability to attach CPS to cell wall and cause invasive diseases (Morona, Morona & Paton 2006; Byrne *et al.* 2011).

6.1.3 CpsB Structure

CpsB is a phosphotyrosine phosphatase initially postulated to modulate capsule production in pneumococci (Fig. 6.1). CpsB is a member of the polymerase and histidinol phosphate phosphatase (PHP) family and its crystallized structure was the first tyrosine phosphatase of this class to be characterized in Gram-positive bacteria (Hagelueken *et al.* 2009). Other members of the PHP family includes capsular biosynthesis proteins in *Staphylococcus* and *Bacillus*, and subunits within DNA polymerases in *Helicobacter*, *Chlamydia*, and *B. subtilis* (Aravind & Koonin 1998). CpsB is a 243-aa, 26-kDa protein consisting of 2 triple-stranded β-sheets surrounded by eight α-helix rings (Fig. 6.1).

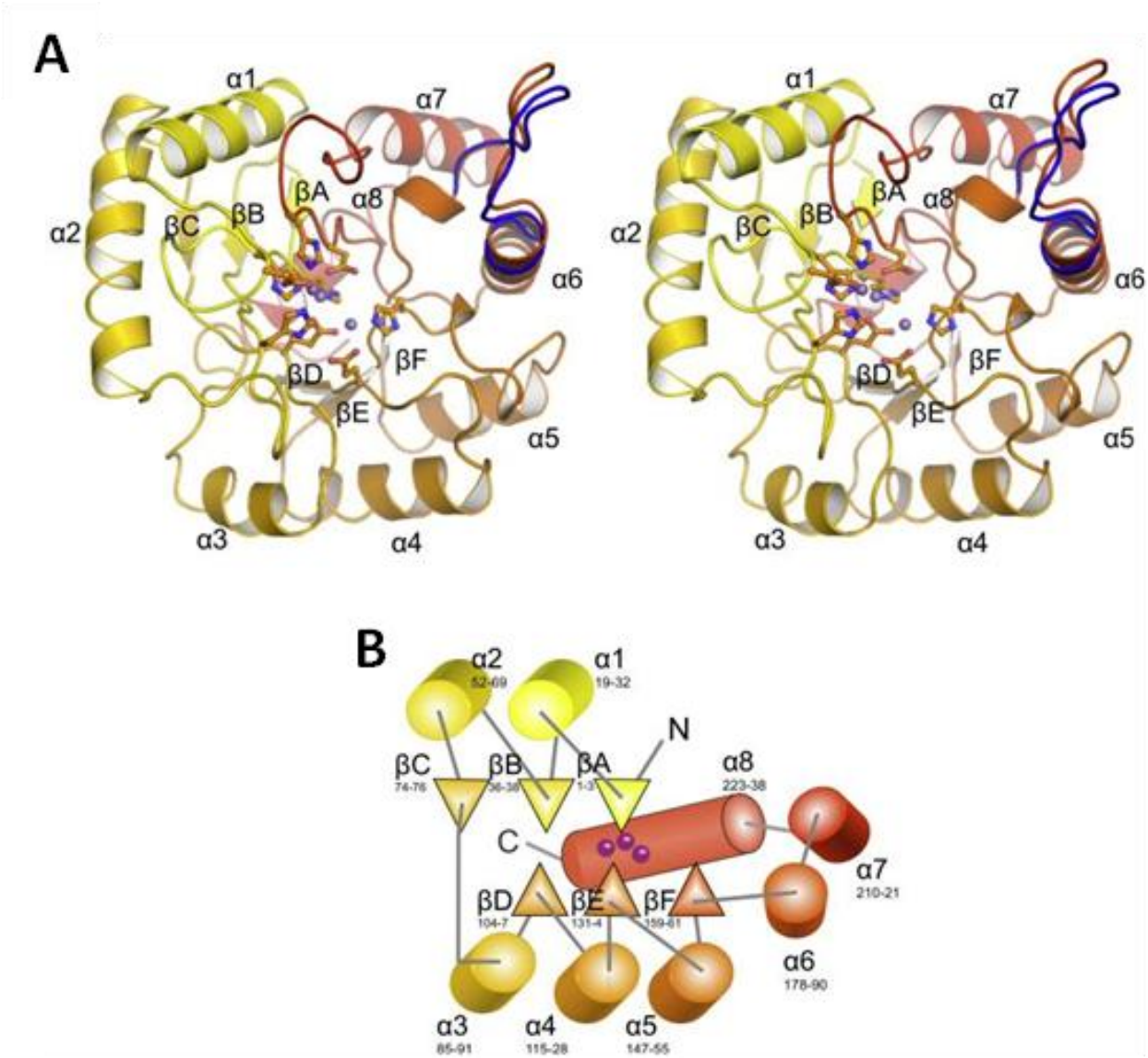


Fig. 6.1 Crystal (A) and schematic (B) structures of CpsB. The phosphotyrosine phosphatase **is** composed of eight α-helices and three pairs of β-sheets in parallel. The blue and purple spheres represent metal ions required for full enzymatic activity. Figure from Hagelueken *et. al.* (Hagelueken *et al.* 2009).

CpsB activity was documented to be dependent on the presence of Mn^{2+} (Morona, 2001), and addition of Mn^{2+}, as well as Cu^{2+} and Co^{2+} ions was associated with up to 10-fold increase in biological activity (Hagelueken, 2009). The active site contains 3 metal-binding pockets, which are thought to involve two arginine residues at positions 139 and 206. When the enzyme is bound to phosphate, there is minimal conformational change to the enzyme, except that the metal-binding pockets are replaced by the phosphate molecule. Mutants with arginine-to-alanine substitutions to position 139 have 95% reduction in CpsB enzymatic activity, suggesting that this

residue is crucial for enzymatic function (Hagelueken, 2009). CpsB also contains additional conserved motifs containing histidine residues which may be involved in stabilization of the metal cations, and mutation of these residues led to increased levels of phosphorylated forms of CpsD, associated with a rough phenotype and minimal visual differences in morphologies from that of a *cpsB*-deletion Rx1-derivative mutant (Morona, 2001). Alignment of homologues of CpsB with other Gram-positive species revealed that these motifs, in particular histidine residues within these motifs, are conserved (Fig. 6.2) (Morona, 2002). In addition, a series of amino acid residues throughout the protein has been thought to line the phosphate pocket (Kim, 2011). CpsB has 30% amino acid homology to PHP protein YwqE of *Bacillus subtilis*, which is a phosphotyrosine phosphatase likely to be a regulator of exopolysaccharide synthesis (Mijakovic *et al.* 2003).

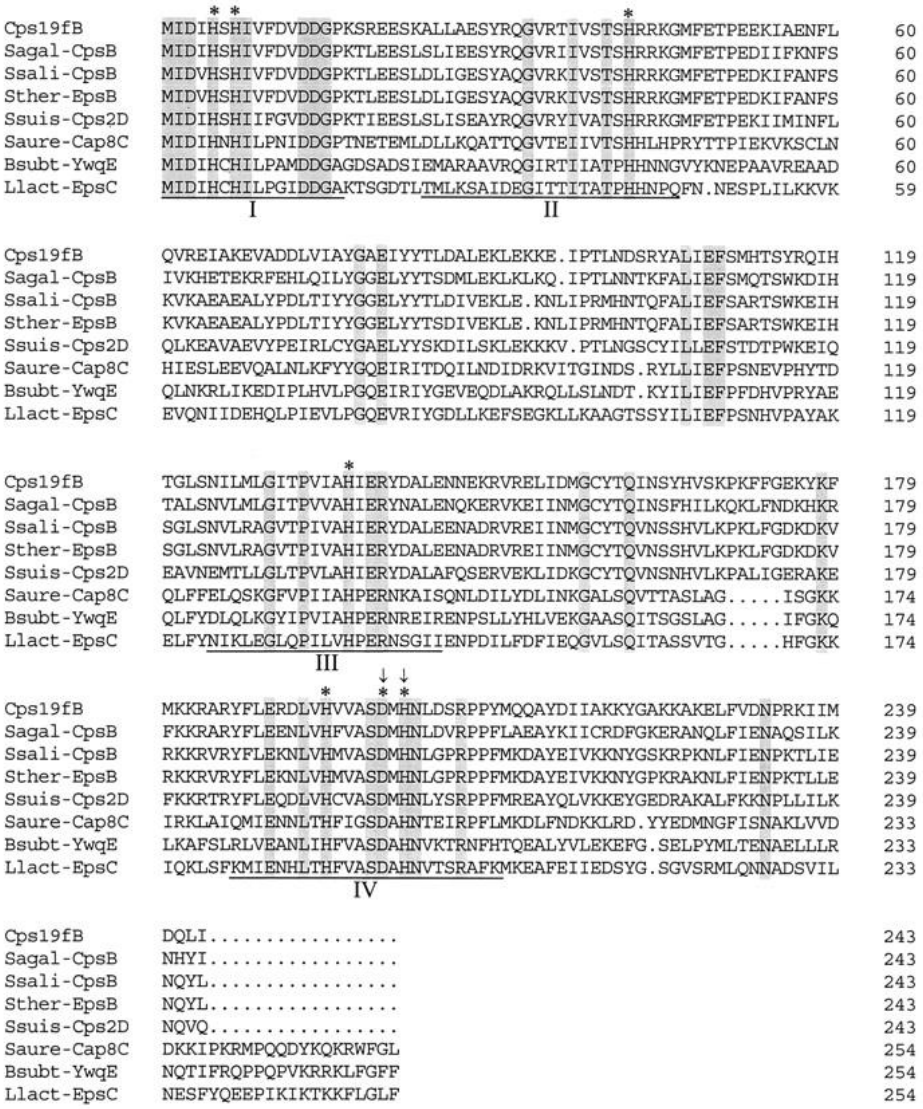


Fig. 6.2 Alignment of *S. pneumoniae* CpsB and its homologues across Gram-positive species. Alignment included members of *S. agalactiae*, *S. salivarius*, *S. thermophilus*, *S. suis*, *S. aureus*, *B. subtilis*, and *Lactococcus lactis*. Members of the PHP superfamily proteins contain four conserved motifs indicated by Roman numerals. Shaded amino acids indicated conserved positions, including residues in asterisks that are thought to be important for catalysis. Figure from Morona *et al.* (Morona *et al.* 2002). Journal of Bacteriology authorizes the reuse of journal material including figure shown for academic purposes

6.1.4 Genetic Characteristics of *cpsB*

Compared to serotype-specific regions of the *cps* locus, sequences of regulatory genes including *cpsB* between serotypes are more conserved in that it is present in all serotypes. Studies investigating *cpsB* sequence variation are scarce. While understanding the genetic organization of *cps* locus of serotype 19F, Morona *et al.* (Morona, Morona & Paton 1997) documented the hybridization of *cpsB* from 18 serotypes, including serotype 19A, to *cps19fB.* However, subsequent sequence analysis between *cps19fB* and *cps19aB* reveals that only a short region within this gene shows 100% identity between the two serotypes, while the overall sequence similarity of *cpsB* between is only 82%, suggesting that sequence analysis reveals a higher degree of variation in this gene (Morona, Morona & Paton 1999a). Analysis of *cps19aB* reveals the presence of mosaic patterns with blocks of GC% variation (Morona, Morona & Paton 1999a); the authors suggest that recombination of short fragments into this gene has occurred (Morona, Morona & Paton 1999a).

Sequence analysis of *cpsB* from 11 serotypes revealed two major classes, with sequence variation between the classes up to 27% (Jiang, Wang & Reeves 2001). Class I includes serotypes 1, 14, 18C, and 19F, while class II consists of serotypes 2, 6B, 8, 19A, 23F, and 33F. In contrast to the high sequence variation between the two classes, sequence variations between serotypes within the same class do not exceed 4%. This indicates that the two clusters are well defined with serotypes within each cluster highly similar.

In agreement to this, a recent study involving a larger number of serotypes revealed the dichotomous divergence of *cpsB* sequences (Varvio *et al.* 2009). In this study, the clusters are classified blue (class I) and red (class II), with a greater genetic diversity within sequences of the blue cluster, suggesting that this cluster is the ancestral group. The lower GC content (lower by 5% compared to blue cluster) in the red cluster suggests that the *cps* regulatory genes of these serotypes originate from an external gene pool of unknown origin. Serotypes within a serogroup may belong to different clusters, indicating that similarity in capsule structure is not correlated with sequence similarity in this region. Analysis of epidemiological data shows that class I serotypes are more likely to associate with carriage while serotypes in class II are predominantly found in IPD (Varvio *et al.* 2009). However, this finding was based

upon a small selected number of cohorts, and the strength of this association has not been determined in other studies. Nonetheless, sequence variation at the *cps* regulatory region may play a role in its ability to colonize and cause disease, perhaps by modulating the production of capsules.

6.1.5 Aim

The discrepancies observed between *in silico* and experimental results in sequetyping resolution power as described in Chapter 3 led us to hypothesize that strains within a serotype may have different *cpsB* sequences. While sequence diversity between representative strains of serotypes has been analyzed, information regarding sequence diversity within a serotype is limited to a minority of serotypes (McEllistrem *et al.* 2004; Varvio *et al.* 2009; Elberse *et al.* 2011). Variation in *cpsB* within a serotype may allow the greater potential to regulate capsule expression under different environmental conditions.

With the availability of a large number of *cpsB* sequences following the works described in Chapter 3, we aimed to compare intra-serotype diversity of *cpsB* in different pneumococcal serotypes. First the *cpsB* phylogenetic tree published by Varvio *et al.* (Varvio *et al.* 2009) was expanded by adding *cpsB* sequences from the newly-identified serotypes 6C and 6D, as well as incorporating additional *cpsB* sequences from results obtained in our sequetyping work (Chapter 3) and previously published works. The relationships of *cpsB* variation and STs and CCs were also analyzed.

6.2 Materials and Methods

6.2.1 *cpsB* Sequence Data for Pneumococcal Strains

Please refer to Appendix Table A1 for strains included in this chapter. Sequence data for an additional 172 *cpsB* sequences were available from the GenBank database. Accession numbers for these Genbank entries are available in Appendix Table A2.

6.2.2 Multilocus Sequence Data

Sequence types (STs) were available for 163 strains in this study. ST data for strains from HPA and SHLMPRL were provided by Dr. Bruno Pichon and Dr. Giles Edwards, respectively. STs of strains with *cpsB* sequences from GenBank are available in respective publications. ST from Tanzanian strains were from works in Chapter 4. MLST was performed on invasive pneumococcal strains from the Royal Free Hospital as described in Chapters 2 and 3. eBURST (Feil *et al.* 2004) was performed to illustrate genetic relatedness of all STs included in this study, and to identify clonal complexes of related STs differing in one of the seven MLST loci (single-locus variants, SLVs).

6.2.3 Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignments on *cpsB*, concatenated MLST sequences (without *ddl*), and translated CpsB amino acid sequences were constructed on MEGA software, version 5.05 (Tamura *et al.* 2011). GC contents were calculated by MEGA. Topologies based on *cpsB* alignments were acquired using the neighbour-joining (NJ) and maximum likelihood (ML) methods, with 1,000 bootstrap repetitions for NJ and 100 bootstrap for ML. NJ analysis, based on matrix of pairwise distance differences, is more computationally efficient and was also used by Varvio *et al.* (Varvio *et al.* 2009), but ML, inferring phylogenetic relationship directly from each site in the multiple alignment, provides the most probable estimate of phylogenetic relationship based on a given evolutionary model. The NJ trees are included in the results section of this chapter, whereas the ML trees of the same alignment data are included in the appendix. Bootstrap values < 75% are not included in the trees.

The ML trees were constructed using the Tamura-3 Paramete evolutionary model. The Tamura 3 parameter model does not assume that each nucleotide substitution occurs at the same rate (i.e. unlike the Jukes and Cantor model). Specifically, branch

lengths in ML under this model are calculated based on two separate calculations: transversion and transition nucleotide substitutions. Also, the Tamura 3 parameter accounts for mutational bias along a nucleotide site on the alignment. It was chosen over other models from the MEGA software as providing the most likely phylogenetic representation based on the sequence alignment data.

6.2.4 Statistical Analysis for Intra-Serotype Diversity

The number of polymorphic or segregating sites (S), haplotype diversity (Hd), nucleotide diversity (π), and their standard deviations were calculated using DnaSP5 software for *cpsB* for strains of serotypes 1, 3, 4, 5, 6A, 6B, 6C, 7F, 8, 9V, 10A, 14, 15A, 17F, 18C, 19A, 19F, 21, and 23F. For all but serotypes 10A, 15A, and 21, the aforementioned parameters were calculated also on concatenated ST data; serotypes 10A, 15A, and 21 did not have enough ST data on more than one strain. Intra-serotype Hd corresponds to the likelihood of two sequences to be of different alleles, whereas intra-serotype π is the average number of substitution per site. The comparison of *cpsB*π between serotypes was calculated by one-way ANOVA and Tukey's test for multiple pairwise comparisons of means. As not all strains with *cpsB* had MLST data, *cpsB*π was repeated for strains with only MLST information. This is to satisfy the calculation of statistical significance between *cpsB*π and MLSTπ using the paired t-test. It is possible that differences in π were due to differences in π of MLST data. Therefore, the ratio of *cpsB*π to MLSTπ was calculated for each serotype to normalize for MLST differences.

6.3 Results

6.3 1 Multiple Alignment of *cpsB* of Different Serotypes

A total of 351 sequences of 92 serotypes were available for analysis. In most serotypes, *cpsB* has a length of 732 bp, encoding for 243 amino acids. Serotype 3 contained a single nucleotide deletion at nucleotide position 282, leading to a frameshift mutation and a premature stop codon at amino acid position 106. The sequence of new serotype 11E (Calix *et al.* 2010) was not available at the time of study. Of all 351 sequences, fifty percent of the sequence (367/732 bp) belongs to polymorphic sites.

6.3.2 Phylogenetic Tree of *cpsB*

A neighbor-joining phylogenetic tree (Fig. 6.3, maximum likelihood tree in Appendix Fig. A2, neighbor-joining radial tree with taxon names in Appendix Fig. A3, and maximum-likelihood radial tree with taxon names in Appendix Fig. A4) based on 351 *cpsB* sequences of 92 serotypes showed a similar two-cluster topology as seen previously (Varvio *et al.* 2009). Three serotypes (serotypes 29, 39, 43 indicated in green) showed sequences resembling non-pneumococcal origins in 3' end of *cpsB* (Varvio *et al.* 2009), and are indicated in green in Fig. 6.3. Purple serotypes correspond to sequences where mosaics of blue and red sequences were detected. Serotypes 3 contains a single-nucleotide deletion in this gene, resulting in a premature stop codon in amino acid position 94, and capsule synthesis in serotype 37 is not dependent on *cps* locus (Llull *et al.* 1999). The gene *cpsB* in serotypes 3 and 37 are not expressed and are indicated in black. Serotypes 25A and 38 formed a very divergent group that did not cluster with either blue or red sequences. However they were considered red serotypes based on the *cpsB* GC content (35%).

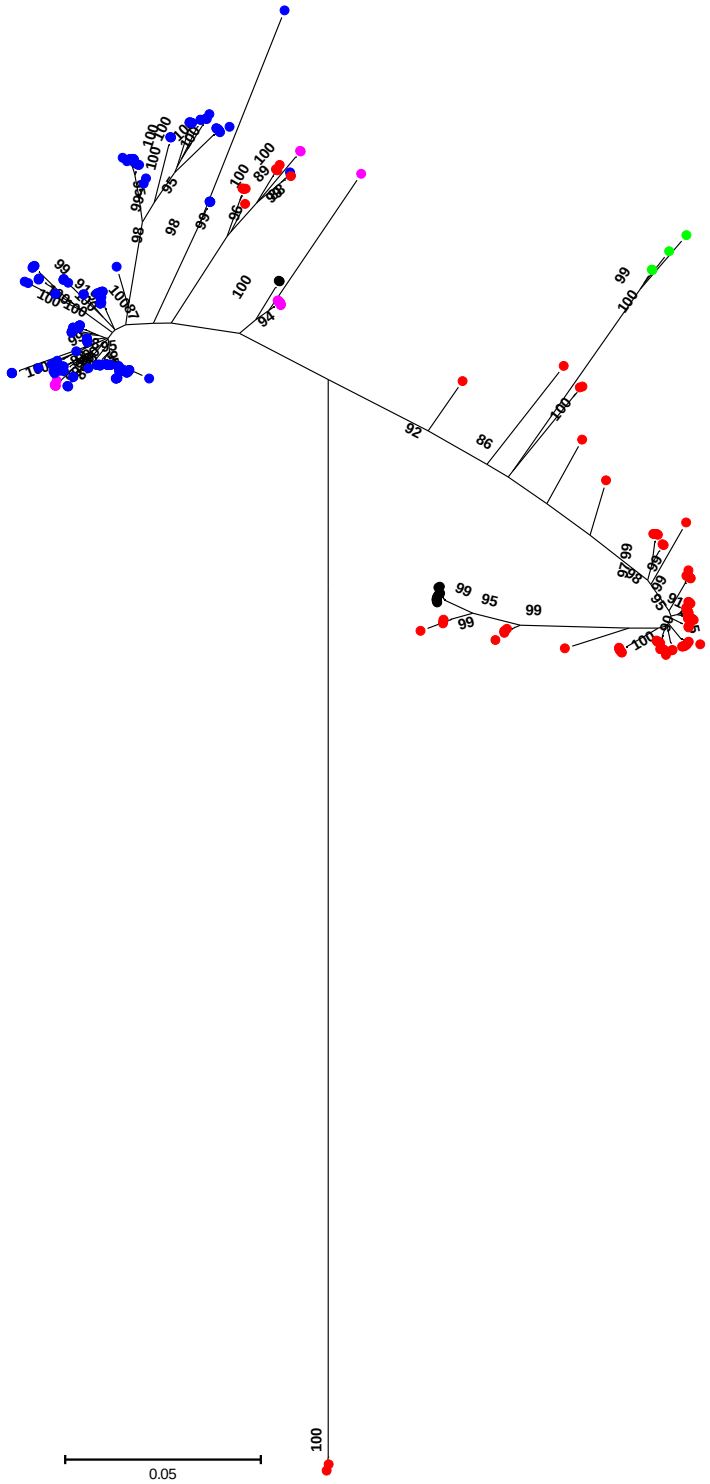


Fig. 6.3. Neighbour-joining phylogenetic representation of pneumococcal *cpsB* sequences. Red and blue colouring was based on study of Varvio *et al.* (Varvio *et al.* 2009), where blue represents serotypes commonly associated with carriage, and red represents those predominantly found in invasive diseases. Serotypes which, according to Varvio *et al.*, had blue-red mosaic within *cpsB* (serotypes 15A, 15F, 23F, 35F, 47A, 47F) are indicated in purple. Serotypes 29, 39, and 43 contained mosaic sequences resembling red sequences and other streptococcal sequences, and are indicated in green. Serotypes 3 and 37 do not require *cpsB* for capsule production, and are indicated in black. Phylogenetic representation performed on MEGA 5.05 with bootstrapping of 1,000 repetitions, with bootstrapping values over 85% indicated. Distances represented as the number of nucleotide substitutions per site. A total of 351 sequences were included in analysis.

Fig. 6.3 shows that despite the higher number of sequences included, the dichotomous topology is preserved from that seen using a single strain per serotype. The presence of *cpsB* of serotype 37 within the blue cluster is expected, as this serotype contains a cryptic *cps33f* locus, a locus in the blue cluster. Serotype 3 strains were grouped into the red cluster, and the lower GC% content (~34%) of this serotype is in agreement with the red/blue division based on this property. Serotypes 6C and 6D strains grouped with some strains of serotype 6B, and only a single nucleotide (T579C) accounted for a difference between 6B and 6C/6D. Serotypes 6C and 6D are identical at *cpsB*.

6.3.3 Intra-Serotype Diversity in *cpsB*

Nineteen serotypes containing at least five strains each were included to compare intra-serotype *cpsB* diversity (Table 6.1). The number of polymorphic sites, allelic variants, GC content variation, haplotype diversity, and nucleotide diversity were tabulated for each of these serotypes. One-way ANOVA revealed that differences exist in both haplotype and nucleotide diversities between serotypes, and four clusters of serotypes can be made based on nucleotide diversity. The highest diversity in cluster one is made up of serotypes 19F, 6A, and 6B, followed by cluster two with serotypes 14, 21, 19A, 5, 17F, and 3. Cluster 3 with serotypes 1 and 9V show yet lower diversity. Serotypes 10A, 15A, 23F, 4, 6C, 8, 18C, and 7F did not show intra-serotype nucleotide diversity. Serotype 6B was significantly more diverse than all serotypes tested, and 6A was also more diverse than other serotypes except 5 and 17F (Tukey's Test, $p < 0.05$).

Table 6.1 Allele and sequence diversity characteristics of *cpsB* and concatenated MLST sequences for individual serotypes

Serotype[α,d,e]	N	S	H	GC% (SD)	*cpsB* Hd[b] (SD)	*cpsBπ*[c] (SD)	No. MLST[d]	MLST S	No. ST	MLST Hd[b] (SD)	MLSTπ[c] (SD)	cpsBπ/MLSTπ
6B	**31**	**54**	**7**	**41.7 (0.4)**	**0.70 (0.059)**	**0.032 (0.0033)**	**20**	**84**	**12**	**0.91 (0.053)**	**0.0088 (0.00094)**	**3.59**
6A	24	12	5	41.8 (0.2)	0.73 (0.059)	0.0045 (0.00058)	22	44	8	0.81 (0.070)	0.0046 (0.00062)	0.98
5	8	7	2	37.7 (0.1)	0.43 (0.17)	0.0041 (0.0016)	5	34	3	0.70 (0.22)	0.0049 (0.0026)	0.84
17F	5	7	2	42.3 (0.2)	0.40 (0.24)	0.0039 (0.0023)	4	18	3	0.83 (0.22)	0.0033 (0.0015)	1.17
19A	**22**	**11**	**7**	**37.7 (0.1)**	**0.54 (0.13)**	**0.0024 (0.00101)**	**17**	**137**	**14**	**0.97 (0.032)**	**0.013 (0.0015)**	**0.18**
19F	**26**	**27**	**6**	**37.7 (0.1)**	**0.76 (0.059)**	**0.0018 (0.00044)**	**14**	**106**	**13**	**0.99 (0.031)**	**0.012 (0.0014)**	**0.15**
21	5	2	2	41.4 (0.1)	0.60 (0.18)	0.0016 (0.00048)	-	-	-	-	-	-
14	12	3	4	37.4 (0.1)	0.64 (0.13)	0.0011 (0.00030)	7	8	2	0.48 (0.17)	0.0014 (0.00050)	0.86
1	**13**	**2**	**2**	**37.7 (0.1)**	**0.15 (0.13)**	**0.00042 (0.00034)**	**12**	**34**	**7**	**0.894 (0.063)**	**0.0062 (0.00065)**	**0.067**
3	13	2	3	37.7 (0.1)	0.30 (0.16)	0.00042 (0.00023)	11	0	1	0 (0)	0 (0)	-
9V	14	1	2	42.5 (0.1)	0.14 (0.12)	0.0002 (0.00016)	6	0	1	0 (0)	0 (0)	-
10A	5	0	1	42.3 (0)	0 (0)	0 (0)	-	-	-	-	-	-
15A	5	0	1	42.1 (0)	0 (0)	0 (0)	-	-	-	-	-	-
23F	13	0	1	40.7 (0)	0 (0)	0 (0)	9	44	7	0.92 (0.092)	0.0047 (0.0016)	N/A
4	12	0	1	41.9 (0)	0 (0)	0 (0)	7	8	2	0.48 (0.17)	0.0014 (0.00050)	N/A
6C	5	0	1	42.3 (0)	0 (0)	0 (0)	4	34	2	0.50 (0.27)	0.0062 (0.0033)	N/A
8	7	0	1	41.4 (0)	0 (0)	0 (0)	3	0	1	0 (0)	0 (0)	-
7F	8	0	1	37.2 (0)	0 (0)	0 (0)	6	0	1	0 (0)	0 (0)	-
18C	7	1	2	37.3 (0)	0 (0)	0 (0)	6	0	1	0 (0)	0 (0)	-

[a] N = number of strains, S = Polymorphic sites, H = Number of alleles

[b] Hd = Haplotype diversity: the likelihood of two strains having different *cpsB* alleles or ST

[c] π = Nucleotide diversity: the average number of nucleotide substitutions per site

[d] Number of strains with MLST information. MLST was not available for all strains included for *cpsB* analysis. Serotypes 21, 10A, 15A, and 18C did not have enough MLST data

[e] Serotypes with statistically significant variations between *cpsB* sequences and MLST concatenated sequences by paired t-test ($p<0.05$) are indicated in bold. Variation of *cpsB* for these strains calculated only for strains with MLST data

The diversity of *cpsB* within serotypes was compared with that of concatenated MLST sequences for strains of known STs. This was performed on eight serotypes, as other serotypes did not have sufficient MLST data. ST differences between strains of the same serotype were normalized by dividing *cpsB* π to MLST π. This ratio of *cpsB* nucleotide diversity to MLST diversity is thus a representation of the diversity of this gene relative to that of the genetic background. A high ratio would represent diverse *cpsB* given a particular collection of STs. Analysis shows that serotype 6B has the highest ratio (3.59), while serotype 1 has the lowest (0.067), with a difference of greater than 50-fold.

6.3.4 Relationship between *cpsB*, Clonal Complexes (CCs) and STs

To understand *cpsB* diversity within CCs and STs, a collection of 163 strains with MLST data was represented as an eBURST diagram (Fig. 6.4). A total of 89 STs were identified, thirty-one of which constituted 14 CCs of related strains.

6.3.4.1 Diversity of *cpsB* within CCs of Same Serotypes

Ten of the 14 CCs contained STs of the same serotype, and all but one of these had identical *cpsB* sequences within the CC. CC with ST4429 and 4432 shared one polymorphic site.

6.3.4.2 Diversity of *cpsB* within STs of Same Serotypes

Five STs with more than five strains each were included to assess for *cpsB* diversity within a ST of the same serotype. STs 65, 176, 180, 191, and 246 were associated with serotypes 6A, 6B, 3, 7F, and 4, respectively. Strains within each STs showed identical *cpsB* sequences, except for members of ST180, which showed a single polymorphic site between 11 strains.

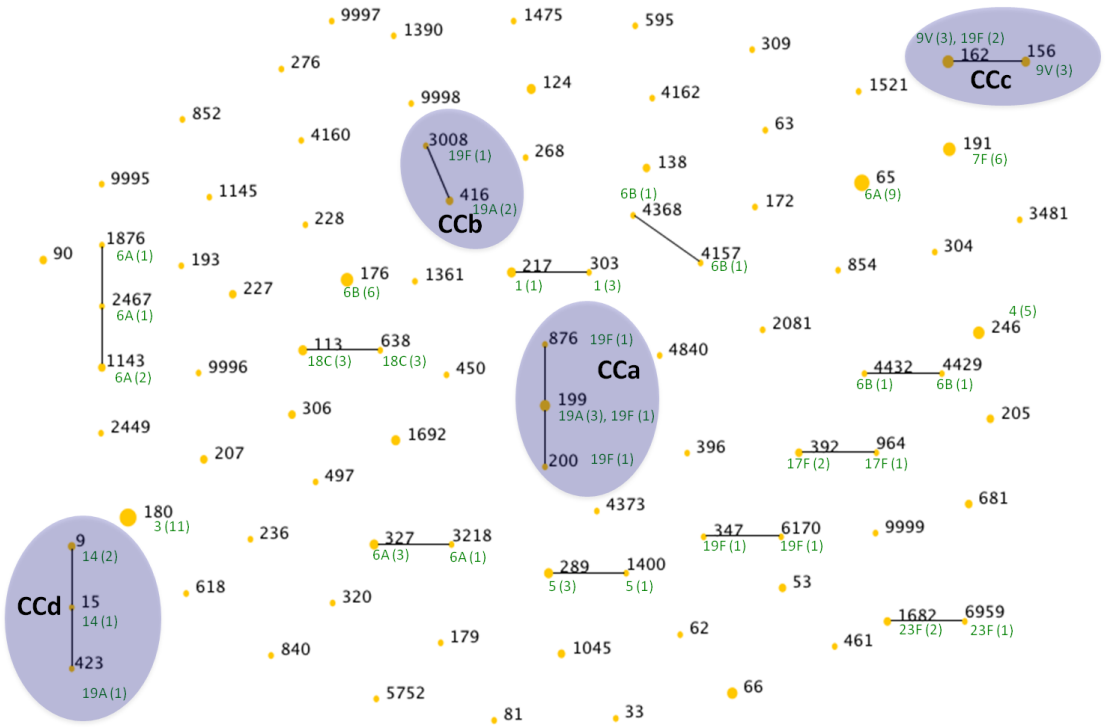


Fig. 6.4. eBURST representation of genetic relatedness between strains with known STs in this study. Single-locus variants (SLVs) differing in only one of the seven MLST loci are connected by solid lines, forming a clonal complex (CC) of related strains. A total of 14 CCs was included in this study. For STs included in analysis of intra-CC and intra-ST *cpsB* diversity, the associated serotypes for each ST are indicated in green, followed by the number of strains in parenthesis. Four of the 14 CCs (CCa, CCb, CCc, CCd) contained different serotypes, and the *cpsB* diversity for these CCs are presented in Table 6.2.

6.3.4.3 Diversity of *cpsB* within CCs and STs of Different Serotypes

Four of the 14 CCs contained related strains of different serotypes. Two clonal complexes each constituted of serotypes 19A and 19F (CCa and CCb), one CC contained members of 9V and 19F (CCc), and one contained STs of 14 and 19A (CCd). All four CCs showed extensive polymorphic sites, but mostly when strains of different serotypes were compared (Table 6.2). STs of the same serotype within each CC still showed identical *cpsB*, except for serotype 19A members of ST199, showing two polymorphisms. Both ST199 and ST162 contained strains of different serotypes, and in both cases polymorphisms were seen only when the different serotypes were compared.

Table 6.2 Diversity of *cpsB* within CCs and STs containing different serotypes

CC	S in CC[a]	ST	Serotype (N)	Intra-ST S
CCa	134	199	19A (3), 19F (1)	134 between serotypes, 2 within 19A
		200	19F (1)	-
		876	19F (1)	-
CCb	111	416	19A (2)	None within 19A
		3008	19F (1)	-
CCc	120	162	9V (3), 19F (2)	120 between serotypes, none within serotypes
		156	9V (3)	0
CCd	136	9	14 (2)	136 between serotypes, none in serotype 14
		15	14 (1)	-
		423	19F (1)	-

[a] S: segregation/polymorphic sites

6.3.5 Relatedness of *cpsB* within and between Serotypes

A NJ tree was reconstructed including only serotypes with multiple strains to compare *cpsB* diversity within and between serotypes (Fig. 6.5). A ML tree of the same data is provided in Appendix Fig. A5. In agreement to the variation seen in haplotype and nucleotide diversities, serotypes 6B and 6A contained strains distantly related in terms of *cpsB* sequences. Serotype 6A members grouped into two major clusters and a single heterogeneous strain. For serotype 6B, most strains belonged to two divergent clusters, one associating with serotype 6A (cluster I), while another clustered closer to members of serotype 19A (cluster II). This divergent cluster is 97% identical in *cpsB* to serotype 11A strains, while the identity between the two divergent 6B clusters is 94%.

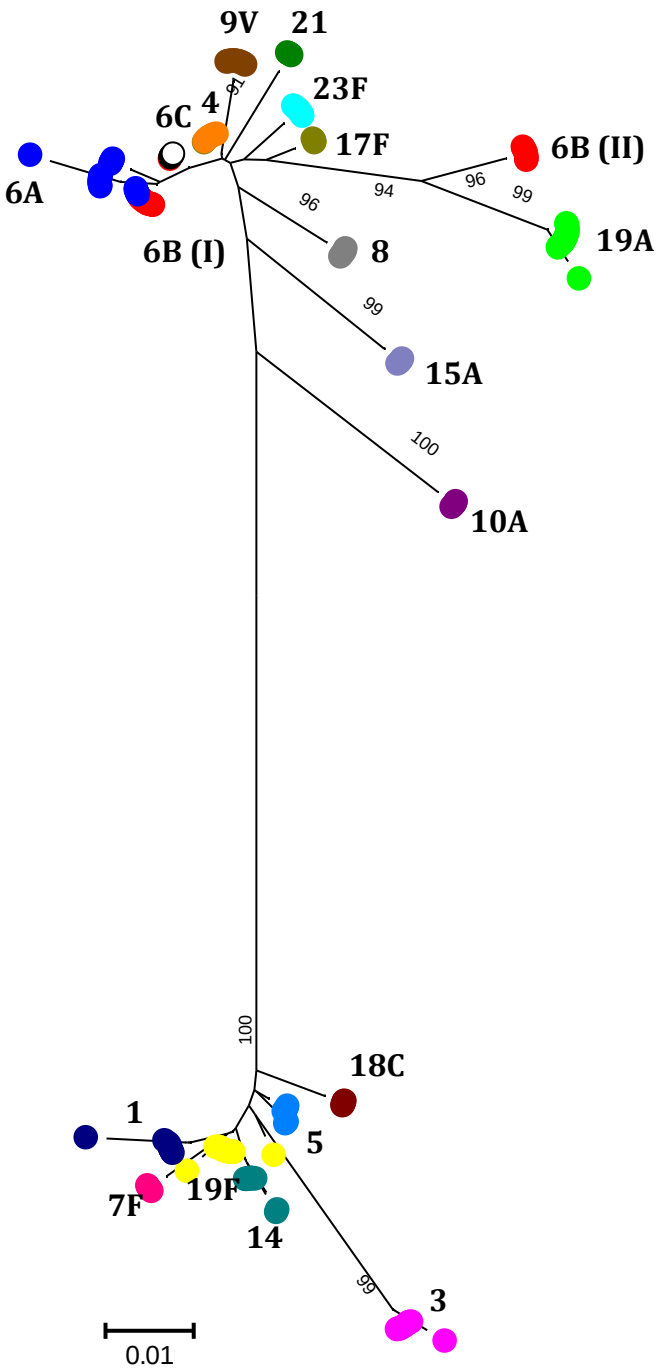


Fig. 6.5. Neighbor-joining phylogenetic representation of serotypes with more than five strains each, colour-coded by serotypes. Serotypes are indicated beside clusters of taxa. Bootstrap with replacements was performed with 1,000 repetitions, and bootstrap values above 85% ware indicated. Taxa are colour-coded based on serotype for ease of visualization. Ruler indicates average number of nucleotide differences per site. Phylogenetic analysis performed on MEGA 5.05.

6.3.6 Locations of Amino Acid Substitutions Between Serotypes

The nucleotide polymorphisms detected across all pneumococcal serotypes corresponded to 116 amino acid substitutions (47.7% of 243 amino acids in CpsB). The locations of these substitutions were determined in reference to known amino acids to form secondary structures and those required for catalysis (Morona *et al.* 2002; Hagelueken *et al.* 2009). For this analysis, strains of serotypes 3 and 37 were not included, as functional expression of CpsB is not required for these serotypes. A plot of substitution along the length of the protein revealed that most non-conservative substitutions occurred within domains α4 and α7 (Fig. 6.6A). For the amino acids thought to be involved in metal- and phosphate-binding (His5, His7, Asp14, His42, Glu80, Glu108, His136, Arg139, His194, Asp199, His201, Arg206), all but His194 and His201 were conserved in all sequences. Position 194 was substituted with a tyrosine residue in a single serotype 19F strain, and His201 was substituted with tyrosine in all serotype 6A strains and the serotype 6B strains closely related to 6A.

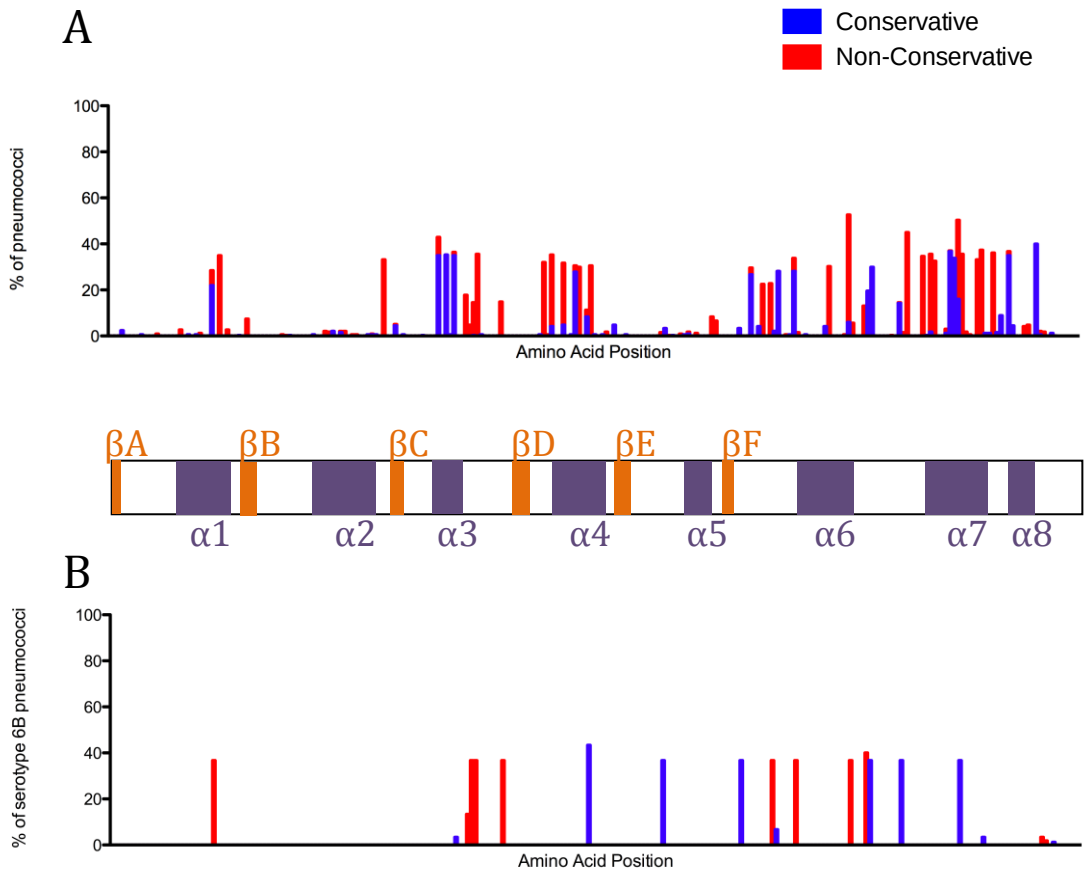


Fig. 6.6. Amino acid substitutions along CpsB for all pneumococcal serotypes except serotypes 3 and 37 (A) and strains within serotype 6B (B). For both histograms, the x-axis refers to amino acid position (total of 243 positions), and the y-axis refers to proportion in percentage of sequences with amino acid substitutions. Blue bars represent conservative substitutions that do not change the polarity and acidity of the amino acid, and red bars represent non-conservative substitutions leading to a change in polarity and/or acidity of the amino acid. Locations of globular domains are indicated as α-helices (purple) amd β-sheets (orange). A total of 338 sequences were included in Fig. 6.6A and 30 in Fig. 6.6B.

6.3.7 Serotype Distributions of α4 and α7 Non-Conservative Substitutions

Detailed alignments of the α4 (Fig. 6.7A) and α7 (Fig. 6.7B) regions reveal that the majority of non-conservative substitutions were detected within the red serotypes, with the most common substitutions in positions 112, 114, 117, 121, and 124 in α4 and positions 211, 212, 218, and 219 in α7. The purple serotypes with mosaics of blue and red also show substitutions in those regions.

A

Serotype	Amino Acid Position																
	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
2 D39	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	S
6A 11 3085	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
6A 11 5346	N	T	P	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
6B 11NP10Jan	N	T	P	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
6B 2616/39	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
6C 52520075	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
6D MNZ920	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
8 573/62	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
9A Wilder	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
9V N98	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
10A 10061/38	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
10B 423/82	D	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
10C Gro Norge	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
10F SG10F	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
11A 1813/39	N	T	P	Y	R	D	I	H	S	A	L	I	K	I	L	M	L
11B 8087/40	N	T	P	Y	R	D	I	R	S	A	L	S	K	I	L	M	L
11C Eddy nr. 53	N	T	P	Y	R	D	I	R	S	A	L	S	K	I	L	M	L
11D 70/86	N	T	P	Y	R	D	I	H	S	A	L	I	K	I	L	M	L
11F 34356	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
12A 559/66	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
12B SG12B	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
12F 6312	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
17A nr.4704	H	T	S	Y	R	E	I	H	S	A	L	S	K	I	L	M	L
17F Rose	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
18F Gethens	N	T	P	Y	R	D	I	H	S	A	L	I	K	I	L	M	L
19A TCH8431	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
21 N92	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
22A 3405/39	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
22F SG22F	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
23B 1031/41	N	T	P	Y	R	D	I	H	S	A	L	I	K	I	L	M	L
23B 7OP9	N	T	P	Y	R	D	I	H	S	A	L	I	K	I	L	M	S
23B SG23B	N	T	P	Y	R	D	I	H	S	A	L	I	K	I	L	M	L
24B 2236/42	N	T	L	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
24F SG24f	N	T	L	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
25F 601/62	N	T	L	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
27 84320293	N	T	L	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
28A 1982/45	N	T	L	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
28F 34372	N	T	L	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
31 nr.34374	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
32A 2813/41	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
32F nr.34375	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
33A Biehl	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
33B SG33b	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
33C SG33c	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
33D SG33d	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
33F SG33fb	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
34 676/74	H	T	S	Y	R	E	I	H	S	A	L	S	K	I	L	M	L
35A SG35A	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
35B SG35B	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
35C 7765/43	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
36 SG36	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
37 7077-39	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
41A 6803	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	S
41F 8211/40	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
42 198/71	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
44 Hammer	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
46 Eddy nr.73	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
1 3OP2	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
4 WCH35	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
5 Ambrose	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
7A 2040/37	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
7B Johnson	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
7C 1003-2	N	T	L	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
7F SG07f	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
9L T9233/128/68	H	T	S	Y	R	E	I	H	A	G	L	S	N	I	L	M	L
9N 533/62	H	T	S	Y	R	E	I	H	A	G	L	S	N	I	L	M	L
13 34357	H	T	S	Y	R	E	I	H	A	G	L	S	N	I	L	M	L
14 10 1688	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
15B 7904/39	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
15C 533/62	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
16A R105	H	T	S	Y	R	E	I	H	T	G	L	S	N	I	L	M	L
16F nr.34361	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
18A 8609/43	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
18B 1033/41	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
18C WCH94	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
19B nr.4594	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
19C 7588/39	N	T	L	Y	R	D	I	H	N	A	L	S	K	I	L	M	L
19F SPVA96	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
20 34365	H	T	S	Y	R	E	I	H	A	G	L	S	N	I	L	M	L
23A 1196/45	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
24A 2748/40	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
24F L	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
36 1095/39	H	T	S	Y	R	E	I	H	T	G	L	S	N	I	L	M	L
40 Colemore	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
45 Eddy nr.72	N	T	P	Y	R	D	I	H	N	A	L	S	K	I	L	M	L
58 26/63	H	T	S	Y	R	E	I	H	A	G	L	S	N	I	L	M	L
15A 389/39	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
15F 688/63	N	T	P	Y	R	D	I	H	S	A	L	N	K	I	L	M	L
23F UK-577	N	T	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
29 nr.34373	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
35F 361/39	N	I	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
39 203/40	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
43 2427/48	H	T	S	Y	R	E	I	H	T	G	L	S	N	I	L	M	L
47A L351	H	T	S	Y	R	Q	I	H	T	G	L	S	N	I	L	M	L
47F Eddy nr.52	N	I	P	Y	R	D	I	H	S	A	L	S	K	I	L	M	L
25A tp25/38	E	T	S	Y	R	S	I	H	S	A	L	Q	Q	L	L	S	G
38 9687/39	E	T	S	Y	R	S	I	H	S	A	L	Q	Q	L	L	S	G

B

Serotypes	Amino Acid Position											
	210	211	212	213	214	215	216	217	218	219	220	221
2 D39	M	A	E	A	Y	D	L	V	T	Q	K	Y
6A-KEE6A-Vb	M	A	E	A	Y	D	L	V	T	Q	K	Y
6B 11NP10Jan	M	A	E	A	Y	D	L	V	S	Q	K	Y
6B 2616/39	M	A	E	A	Y	D	L	V	T	Q	K	Y
6C-KEE6C-Ic	M	A	E	A	Y	D	L	V	T	Q	K	Y
6D MNZ21	M	A	E	A	Y	D	L	V	T	Q	K	Y
8 573/62	M	A	E	A	Y	D	F	V	T	Q	K	Y
9A Wilder	M	A	E	A	Y	D	L	V	T	Q	K	Y
9V N98	M	A	E	A	Y	D	L	V	T	Q	K	Y
10A SG10A	M	F	E	A	Y	S	I	V	A	K	K	Y
10B 423/82	M	F	E	A	Y	S	I	V	A	K	K	Y
10C Gro Norge	M	A	E	A	Y	D	L	V	S	Q	K	Y
10F SG10F	M	A	E	A	Y	D	L	V	S	Q	K	Y
11A SG11A	M	A	E	A	Y	D	L	V	S	Q	K	Y
11B 8087/40	M	A	E	A	Y	D	L	V	S	Q	K	Y
11C Eddy nr. 53	M	A	E	A	Y	D	L	V	S	Q	K	Y
11D 70/86	M	A	E	A	Y	D	L	V	S	Q	K	Y
11F 34356	M	A	E	A	Y	D	L	V	T	Q	K	Y
12A 559/66	M	A	E	A	Y	D	L	V	T	Q	K	Y
12B Gambia 1/81	M	A	E	A	Y	D	L	V	T	Q	K	Y
12F SG12f	M	A	E	A	Y	D	L	V	T	Q	K	Y
17A nr.4704	M	A	E	A	Y	D	L	V	T	Q	K	Y
17F Rose	M	A	E	A	Y	D	L	V	T	Q	K	Y
18F Gethens	M	A	E	A	Y	D	L	V	S	Q	K	Y
19A 12 1579	M	A	E	A	Y	D	L	V	S	Q	K	Y
21 N92	M	A	E	A	Y	D	L	V	T	Q	K	Y
22A 3405/39	M	A	E	A	Y	D	L	V	T	Q	K	Y
22F SG22F	M	A	E	A	Y	D	L	V	T	Q	K	Y
23B 1031/41	M	A	E	A	Y	D	L	V	S	Q	K	Y
24B 2236/42	M	Q	Q	A	Y	D	I	I	A	K	K	Y
25F 601/62	M	Q	Q	A	Y	D	I	I	A	K	K	Y
27 84320293	M	A	E	A	Y	D	L	V	T	Q	K	H
28A SG28A	M	A	E	A	Y	D	L	V	T	Q	K	Y
28F 34372	M	A	E	A	Y	D	L	V	T	Q	K	Y
31 SG31	M	A	E	A	Y	D	F	V	T	Q	K	Y
32A 2813/41	M	A	E	A	Y	D	L	V	T	Q	K	Y
32F nr.34375	M	A	E	A	Y	D	L	V	T	Q	K	Y
33A Biehl	M	A	E	A	Y	D	L	V	T	Q	K	Y
33B E294	M	A	E	A	Y	D	L	V	T	Q	K	Y
33C SG33c	M	A	E	A	Y	D	L	V	T	Q	K	Y
33D CSF/79	M	A	E	A	Y	D	L	V	T	Q	K	Y
33F SG33fb	M	A	E	A	Y	D	L	V	T	Q	K	Y
34 676/74	M	A	E	A	Y	D	L	V	T	Q	K	Y
35A SG35A	M	A	E	A	Y	D	L	V	T	Q	K	Y
35B 4356/39	M	A	E	A	Y	D	L	V	T	Q	K	Y
35B SG35B	M	A	E	A	Y	D	L	V	T	Q	K	Y
35C 7765/43	M	A	E	A	Y	D	L	V	T	Q	K	Y
41A 6803	M	A	E	A	Y	D	L	V	T	Q	K	Y
41F 8211/40	M	A	E	A	Y	D	L	V	T	Q	K	Y
42 198/71	M	A	E	A	Y	D	L	V	T	Q	K	Y
44 Hammer	M	A	E	A	Y	D	L	V	T	Q	K	Y
46 Eddy nr.73	M	A	E	A	Y	D	L	V	T	Q	K	Y
1 3OP2	M	Q	Q	A	Y	D	I	I	A	K	K	Y
4 11 4573	M	A	E	A	Y	D	L	V	T	Q	K	Y
5	M	Q	Q	A	Y	D	I	I	A	K	K	Y
7A 2040/37	M	Q	Q	A	Y	D	I	I	A	K	K	Y
7B Johnson	M	Q	Q	A	Y	D	I	I	A	K	K	Y
7C 1003-2	M	Q	Q	A	Y	D	I	I	A	K	K	Y
7F SG07f	M	Q	Q	A	Y	D	I	I	A	K	K	Y
9L T9233/128/68	M	Q	Q	A	Y	D	I	I	A	K	K	Y
9N 533/62	M	Q	Q	A	Y	D	I	I	A	K	K	Y
13 34357	M	Q	Q	A	Y	D	I	I	A	K	K	Y
14 10 1688	M	Q	Q	A	Y	D	I	I	A	K	K	Y
15B 7904/39	M	Q	Q	A	Y	D	I	I	A	K	K	Y
15C 533/62	M	Q	Q	A	Y	D	I	I	A	K	K	Y
16A R105	M	A	E	A	Y	D	L	V	T	Q	K	Y
16F SG16F	M	Q	Q	A	Y	D	I	I	A	K	K	Y
18A 8609/43	M	Q	Q	A	Y	D	I	I	A	K	K	Y
18B 1033/41	M	Q	Q	A	Y	D	I	I	A	K	K	Y
18C WCH94	M	Q	Q	A	Y	D	I	I	A	K	K	Y
19B nr.4594	M	Q	Q	A	Y	D	I	I	A	K	K	Y
19C 7588/39	M	Q	Q	A	Y	D	I	I	A	K	K	Y
19F 11 6402	M	Q	Q	A	Y	D	I	I	A	K	K	Y
20 34365	M	Q	Q	A	Y	D	I	I	A	K	K	Y
23A 1196/45	M	Q	Q	A	Y	D	I	I	A	K	K	Y
24A 2748/40	M	Q	Q	A	Y	D	I	I	A	K	K	Y
24F L	M	Q	Q	A	Y	D	I	I	A	K	K	Y
36 1095/39	M	A	E	A	Y	D	L	V	T	Q	K	Y
36 SG36	M	Q	Q	A	Y	D	I	I	A	K	K	Y
40 Colemore	M	Q	Q	A	Y	D	I	I	A	K	K	Y
45 Eddy nr.72	M	A	E	A	Y	D	L	V	T	Q	K	Y
48 656/63	M	Q	Q	A	Y	D	T	I	A	K	K	Y
15A 389/39	M	A	E	A	Y	D	L	V	T	Q	K	Y
15F 688/63	M	A	E	A	Y	D	L	V	T	Q	K	Y
23F UK-577	M	A	E	A	Y	D	L	V	T	Q	K	Y
29 nr.34373	M	K	E	A	Y	E	L	I	S	K	Q	Y
35F 361/39	M	Q	Q	A	Y	D	I	I	A	K	K	Y
39 203/40	I	K	E	A	Y	E	L	I	S	K	Q	Y
43 2427/48	M	K	E	A	Y	E	L	I	S	K	Q	Y
47A L351	M	A	E	A	Y	D	L	V	T	Q	K	Y
47F Eddy nr.52	M	Q	Q	A	Y	D	I	I	A	K	K	Y
25A tp25/38	M	A	E	A	Y	E	I	I	S	K	R	Y
38 9687/39	M	A	E	A	Y	E	I	I	S	K	R	Y

Fig. 6.7. Amino acid substitutions in α4 (A) and α7 (B) regions of CpsB. Serotypes listed according to their blue/red designation as defined by Varvio *et al.* (Varvio, 2009). Serotypes in purple represent those with mosaics of blue and red sequences. Serotypes 25A and 38 have divergent sequences but are grouped as red serotypes based on low GC content (~35%). Conservative amino acid substitutions indicated as blue blocks, and non-conservative in red.

6.3.8 Locations of Amino Acid Substitutions Within Serotype 6B

Sequences within serotype 6B showed 19 amino acid substitutions, fourteen of which occurred in the linker regions (Fig. 6.6B and Table 6.3). In the five substitutions within globular domains, only one (20%) had a change in polarity (Arg27Tyr), with the remaining four substitutions with the same polarity and acidity. Changes in polarity and acidity were more common in substitutions occurring in linker regions but this was not significant ($p = 0.30$ Fisher's Exact Test). All but one strain in 6B cluster I was isolated from hosts with pneumococcal diseases, whereas in cluster II both carriage and clinical pneumococci were isolated in similar proportions (Table 6.4). No apparent association between geography and *cpsB* clustering was observed.

Table 6.3. Location and types of amino acid substitutions across CpsB in strains belonging to Clusters I and II of serotype 6B.

Substitution Site	Region	Amino Acids	Acidity and Polarity
27	α1	Threonine Alanine	Neutral, Polar Neutral, Non-Polar
89	α3	Aspartic Acid Glutamic Acid	Acidic, Polar Acidic, Polar
92	L6	Glutamic Acid Glycine	Acidic, Polar Neutral, Non-Polar
93	L6	Lysine Asparagine	Basic, Polar Basic, Polar
94	L6	Lysine Asparagine	Basic, Polar Basic, Polar
101	L6	Aspartic Acid Asparagine	Acidic, Polar Neutral, Polar
123	α4	Serine Asparagine	Neutral, Polar Neutral, Polar
142	L9	Valine Alanine	Neutral, Non-Polar Neutral, Non-Polar
162	L11	Valine Isoleucine	Neutral, Non-Polar Neutral, Non-Polar
170	L11	Proline Serine	Neutral, Non-Polar Neutral, Polar
171	L11	Lysine Arginine	Basic, Polar Basic, Polar
176	L11	Arginine Proline	Basic, Polar Neutral, Non-Polar
190	L12	Glutamine Arginine	Neutral, Polar Basic, Polar
194	L12	Histidine Tyrosine	Basic, Polar Neutral, Polar
195	L12	Valine Isoleucine	Neutral Non-Polar Neutral, Non-Polar

203	L12	Leucine	Neutral Non-Polar
		Valine	Neutral, Non-Polar
218	α7	Threonine	Neutral, Polar
		Serine	Neutral, Polar
224	α8	Alanine	Neutral, Non-Polar
		Glycine	Neutral, Non-Polar
239	L14	Methionine	Neutral, Non-Polar
		Arginine	Neutral, Non-Polar

Table 6.4 Country of isolation of serotype 6B carriage and clinical strains according to clustering by *cpsB* based on amino acid changes as shown on Table 6.3 and invasiveness

Strains	Country of Origin	Carriage/Clinical
Cluster I		
3OP7S	Tanzania	Carriage
71560309	United Kingdom	Clinical
80520052	United Kingdom	Clinical
N252	United Kingdom	Clinical
N380	United Kingdom	Clinical
N465	United Kingdom	Clinical
N6	United Kingdom	Clinical
WCH18	Australia	Clinical
KEE-6B-Ia	Netherlands	Clinical
KEE-6B-Ib	Netherlands	Clinical
KEE-6B-Ic	Netherlands	Clinical
KEE-6B-Id	Netherlands	Clinical
KEE-6B-Ie	Netherlands	Clinical
KEE-6B-If	Netherlands	Clinical
KEE-6B-Ig	Netherlands	Clinical
KEE-6B-IIa	Netherlands	Clinical
KEE-6B-IIb	Netherlands	Clinical
Cluster II		
11NP10Jan	Tanzania	Carriage
35NP1	Tanzania	Carriage
35NP6	Tanzania	Carriage
3NP7Mar	Tanzania	Carriage
670-6B	Spain	Clinical
69OP10	Tanzania	Carriage
N389	United Kingdom	Clinical
N888b	United Kingdom	Clinical
PN94/361	United Kingdom	Clinical
KEE-6B-IIIa	Netherlands	Clinical
KEE-6B-IIIb	Netherlands	Clinical

6.4 Discussion

In Chapter 3, we observed that some strains of a serotype differed in *cpsB* sequences to those of other members of the same serotype deposited onto GenBank. We therefore hypothesized that intra-serotype diversity in this region exists for pneumococci, and that different serotypes have different extents of *cpsB* sequence variations. Regulation of capsule expression is crucial for pneumococcal persistence and survival in different host environments; full encapsulation appears to be beneficial for systemic virulence, while reduced expression of capsular polysaccharide exposes surface proteins facilitating adherence and biofilm formation (Magee & Yother 2001; Hammerschmidt *et al.* 2005; Moscoso, García & López 2006). As *cpsB* has been shown to modulate pneumococcal capsule synthesis (Morona *et al.* 2000; Morona *et al.* 2002; Morona, Morona & Paton 2006), sequence variation in this region within a serotype may result in different strains of a serotype to regulate capsule biosynthesis differently, thereby facilitating its success of particular serotypes in different environments. While sequence variation in capsular regulatory region between serotypes is well known (Morona, Morona & Paton 1997; Morona, Morona & Paton 1999a; Morona, Morona & Paton 1999b; Jiang, Wang & Reeves 2001; Varvio *et al.* 2009), it was not until recently that information regarding intra-serotype *cps* sequence diversity of a handful of serotypes is revealed (Mavroidi *et al.* 2004; Bratcher *et al.* 2011; Elberse *et al.* 2011). Most of these reports document variation at serotype-specific genes. This capsular regulatory gene was selected in this study following our assessment of this region as a target for a single-PCR sequence-based serotyping method (Chapter 3).

6.4.1 Extension of Previous Work

Varvio *et al.* (Varvio *et al.* 2009) included representative strains from pneumococcal serotypes (total 90 strains, one per serotype) and observed a dichotomous phylogeny for *cpsB*, *cpsC*, and *cpsD* sequences, and this pattern was associated with GC content and invasive potential. In this study, by incorporating over 300 *cpsB* sequences covering 92 serotypes, we observe that, while the topology is conserved and GC content does not differ extensively within serotypes, intra-serotype sequence variations were detected. In terms of evolution, the sub-clustering of serotypes suggests the potential for serotypes to undergo sequence variations.

6.4.2 Comparison of Diversity Between Serotypes

Serotypes with multiple *cpsB* sequences can be compared to detected differences in extent of *cpsB* variation. We demonstrated that some serotypes (6B and 6A) exhibit greater diversity than others (1). Serotype 6B, a serotype common in both disease and carriage (Sandgren *et al.* 2004), has been documented to be genetically diverse (Robinson *et al.* 2002). In contrast, up to 90% of serotype 1 strains around five continents are associated with six STs (Brueggemann & Spratt 2003). The observed *cpsB* variation may thus be a reflection of the strain differences.

Therefore, to account for sequence variations in the core genes, *cpsB* diversity was normalized against the core genome (as inferred by concatenated MLST sequences) differences. The ratio of *cpsB* to core gene diversity was shown to differ by at least 50-fold between serotypes 6B and 1, suggesting that these serotypes are undergoing different rates of *cpsB* diversification relative to their core genomes. Phylogenetic analysis reveals that strains of serotype 6B are sub-clustered into two groups, which are more variable to each other than to members of other serotypes. It may be that such differences in *cpsB* allow serotypes to regulate capsule expression to persist in different host environments, and this is in agreement to serotype 6B and 1 being common and rarely detected in carriage, respectively. However, serotypes 19F and 19A, also common in carriage and invasion, showed low rates of *cpsB* diversification compared to the core genes. Of note, serotypes 1 and 19F, indicated as having low intra-serotype *cpsB* variation, have been demonstrated to be relatively resistant to C3 deposition and opsonophagocytosis compared to serotype 6B, (Melin *et al.* 2009; Melin *et al.* 2010).

Intriguingly, others have documented the dichotomous division of serotype 6B pneumococci based on sequences of serotype-specific genes (Mavroidi *et al.* 2004; Bratcher *et al.* 2011). Subsequently, extensive variation in *cps* genes, including those at the 5' end, had been postulated to be evidence of diversification to new serotypes, possibly mediated by recombination events (Elberse *et al.* 2011; Song, Baek & Ko 2011). As stated by Elberse *et al.* (Elberse *et al.* 2011), the divergent serotype 6B may represent a new serotype that cross-react with 6B-specific antisera. Sequencing of these serotype-specific genes in our serotype 6B collection and characterization of capsule structure by NMR would enable comparison between diversity between the

capsule regulatory and synthesis regions, and how these variations relate to structural differences.

6.4.3 Comparisons of Diversity Between CCs

Analyzing *cpsB* differences within CCs revealed that minimal polymorphisms are present within a CC if all clones of that CC are of a serotype. However, when STs of multiple serotypes were present within a CC, extensive polymorphisms were seen between the serotypes. This indicates that *cpsB* variation is more likely to be governed by the serotype than the STs. As it is inferred that STs with different serotypes occurred via capsular switching (Enright & Spratt 1998; Jefferies *et al.* 2004; Hanage *et al.* 2011), the observation suggests that *cpsB* sequences are closely linked to capsular phenotype, and *cpsB* is invariably included in switching events. This is in agreement to previous reports of capsular switches with known recombination breakpoints (Coffey *et al.* 1999; Brueggemann *et al.* 2007). The observed association between *cpsB* sequences and antigenic properties in this study therefore indicates that the sequetyping method targeting *cpsB* described in Chapter 3 can accurately identify serotypes in most capsular switch recombinants. However, analysis of additional sequences are required to validate such claims.

6.4.4 Amino Acid Substitutions in CpsB within and between Serotypes

Polymorphisms across serotypes revealed that these alterations encode amino acid changes, the majority of which occurring in α-helices, notably α4 and α7. The substitutions appear to be associated with the blue-red clustering pattern, as most red serotypes contained non-conservative changes within these two helices. Based on available crystal structures of CpsB (Hagelueken *et al.* 2009; Kim *et al.* 2011), these helices are situated on the periphery of the enzyme. Most amino acids crucial for catalysis are conserved. Therefore, while enzymatic activity of CpsB is confined and evolutionarily conserved, different serotypes possess amino acid variations in other structural regions of the enzyme. Analysis of sequence variation within serotype 6B showed that linker regions constituted the majority of substitutions. While these changes may have no effect on enzymatic activity, may be subjected to constant changes that allow the protein to evolve new structures, confirmations, and additional functions (Tokuriki & Tawfik 2009). Whether these changes lead to differential CpsB folding between serotypes is unknown, and structural characterization of these

different sequences may be helpful. In addition, whether these changes affect enzymatic activity is also not clear; differences in enzymatic activity between these different alleles can be determined by phosphatase assays.

6.4.5 Additional Factors Affecting Capsule Production

Despite the variation in *cpsB* observed between serotypes, capsule production regulation is influenced by other factors. The upstream gene *cpsA* has been shown to bind to the *cps* promoter and expression of downstream genes and affect capsule expression (Morona *et al.* 2000; Hanson, Lowe & Neely 2011). Although no transcriptional control elements have been identified upstream of *cps* promoter, the presence of RUP (Repeat Unit for Pneumococcus) elements upstream of *cpsA* in some serotypes and *tts* in serotype 37 may act as binding sites for regulatory proteins (Hoskins *et al.* 2001; Llull, García & López 2001; Moscoso & García 2009). Sequence duplications within serotype-specific genes (*cps3D*, *cps8E* and *tts* for serotypes 3, 8 and 37, respectively) also contribute to reversible phase variations (Waite, Struthers & Dowson 2001; Waite *et al.* 2003). In addition, transcription of capsule genes or production of capsules can also be altered by genes outside the capsulation locus, such as the carbohydrate pyrophosphorylase gene *galU* (Mollerach, López & García 1998), phosphoglucomutase gene *pgm* (Hardy *et al.* 2001) and the carbohydrate catabolic gene *regM* (Giammarinaro & Paton 2002). Environmental factors also influence *cps* expression; reduction in CPS production has been shown for opaque variants in the presence of oxygen (Weiser *et al.* 2001).

6.5 Concluding Remarks

Comparison of a large international collection of pneumococcal *cpsB* sequences revealed that serotypes differ in the extent of diversity in this gene important for capsule regulation, with serotype 6B being most diverse, up to 50-fold to that of serotype 1. Within strains of 6B two divergent clusters were found. Combining MLST data, we provided evidence that *cpsB* sequences are closely associated with the phenotype of the strain, even in cases of serotype switching. Amino acids required for enzymatic activity remained conserved, while globular domains, especially α4 and α7, showed substitution hotspots. However, whether these non-conservative substitutions contribute to the enzyme structure, activity, or capsule production remains unknown.

CHAPTER SEVEN: Competence Stimulatory Peptide Gene *comC* as a Target for Differentiating *Streptococcus pseudopneumoniae* from *Streptococcus pneumoniae* and Other Streptococcal Species

7.1 Introduction

7.1.1 *Streptococcus pseudopneumoniae*

Streptococcus pseudopneumoniae (the pseudopneumococcus) was recently classified as a species distinct from *S. pneumoniae* (Arbique *et al.* 2004). The pathogenic potential of the pseudopneumococcus has been demonstrated in a murine model, where all mice injected peritoneally with clinical isolates of the organism succumbed within 36 hours (Harf-Monteil *et al.* 2006). Although its detection is thought to be relatively rare (1 pseudopneumococcal isolate in 120 pneumococcal isolates, (Harf-Monteil *et al.* 2006)), pseudopneumococci have been isolated from sputum samples of patients with histories of COPD exacerbations (Arbique *et al.* 2004; Keith *et al.* 2006; Keith & Murdoch 2008; Shahinas *et al.* 2011; Sistek *et al.* 2011). In addition, *S. pseudopneumoniae* forms biofilms *in vitro* and colonizes the nasopharynx (Moscoso, García & López 2006; Johnston *et al.* 2010; Leegaard *et al.* 2010; Domenech, García & Moscoso 2011). Recent genetic and genomic analysis of this organism reveals presence of antibiotic-tolerance and resistance genes (Cochetti *et al.* 2005; Shahinas *et al.* 2011). In addition, non-susceptibilities to a number of antibiotics are common (Keith & Murdoch 2008; Simões *et al.* 2010).

7.1.2 Phenotypic Identification of *S. pseudopneumoniae*

Strains of *S. pseudopneumoniae*, along with *S. mitis* and *S. oralis* have often been classified previously as "atypical pneumococci," because of its similarity to *S. pneumoniae*. *S. pseudopneumoniae* is non-capsulated, α-hemolytic, bile-insoluble, and optochin intermediate-resistant in 5% CO_2 and optochin susceptible in ambient O_2 (Arbique *et al.* 2004). On blood agar pseudopneumococcal colonies appear similar to that of noncapsulated pneumococci, with a colony diameter of up to 1 mm (Keith *et al.* 2006). Under the microscope *S. pseudopneumoniae* also appear in pairs and highly similar to appearance to *S. pneumoniae* (Fig. 5.1).

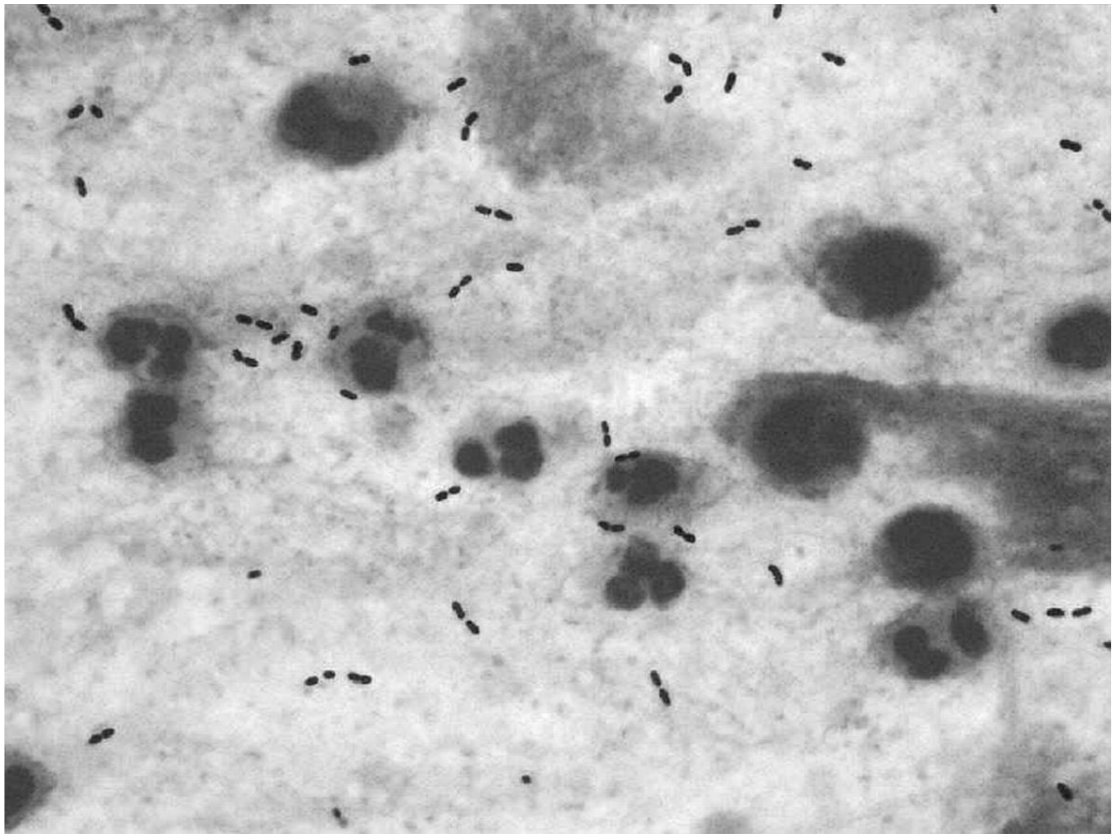


Fig 7.1. *Streptococcus pseudopneumoniae* Gram-stain. Image taken from sputum smear seen under a light microscope at magnification of 100X. Image from Keith *et al*. (Keith *et al.* 2006). Journal of Clinical Microbiology authorizes the reuse of journal material including figure shown for academic purposes.

The pseudopneumococcus reacts positively to the Binax NOW C-polysaccharide-specific antigen test, therefore it cannot be used for its differentiation from the pneumococcus.

7.1.3 Genetic Identification of *S. pseudopneumoniae*

The pseudopneumococcus shares ≥ 99.4% identity to *S. pneumoniae* in 16S rRNA gene sequences (Kawamura *et al.* 1995; Arbique *et al.* 2004 Leegaard *et al.* 2010, Sistek *et al.* 2011). Targeting pneumococcal virulence factors in distinguishing the pseudopneumococcus is challenging, as a large portion of the core pneumococcal genome is detected in *S. pseudopneumoniae* (Johnston *et al.* 2010; Leegaard *et al* 2010; Simões *et al.* 2010). Current genetic approaches attempting to differentiate *S. pseudopneumoniae* from *S. pneumoniae* include targeting *cpsA*, *aliB*-like ORF2 region, *ply*, Spn9802, *lytA*, *sodA*, *recA*, and *psaA*. All these approaches have been proven problematic and require additional analysis (Table 7.1). Sequence analysis of

individual housekeeping genes have been shown to be problematic, and multilocus sequence analysis (MLSA) may reveal divergence between the two species by concatenated gene fragment sequences (Hanage *et al.* 2005; Hoshino, Fujiwara & Kilian 2005; Kilian *et al.* 2008). However, MLSA may not applicable for routine clinical use as the multiple sequencing reactions required for strain identification may be relatively time-consuming and costly.

Table 7.1 Genetic approaches for *S. pseudopneumoniae* identification and differentiation from other Streptococci

Target	References	Comments
rRNA	(Arbique *et al.* 2004; Leegaard *et al.* 2010)	>99% identity with *S. pneumoniae* (Arbique, 2004)
cpsA	(Park *et al.* 2010)	Pseudopneumococcus and noncapsulated pneumococci not differentiable
Spn9802	(Abdeldaim *et al.* 2008; Park *et al.* 2010)	Present in both pneumococcus and pseudopneumococcus Some pneumococci negative for target.
lytA	(Carvalho *et al.* 2007)	Depending on primers used, template targets from some pseudopneumococci may be amplified
aliB-like ORF2	(Simões *et al.* 2011b)	Pseudopneumococcus and noncapsulated pneumococci not differentiable
recA	(Sistek *et al.* 2011)	Clustering of pseudopneumococcal sequences, but may need to be tested on geographically-distinct strains
sodA	(Arbique *et al.* 2004; Hoshino, Fujiwara & Kilian 2005)	Could not be differentiated from *S. pneumoniae*
ply	(Johnston *et al.* 2004; Carvalho *et al.* 2007; Abdeldaim *et al.* 2008)	Present in pseudopneumococci, but may vary in sequence from pneumococcus
psaA	(El Aila *et al.* 2010)	Pseudopneumococcus and pneumococcus differentiable, but may need to be tested on additional strains
MLST/MLVA	(Hoshino, Fujiwara & Kilian 2005; Kilian *et al.* 2008)	Clustering by species based on concatenated sequences, but may be time-consuming for routine clinical diagnosis and identification

7.1.4 Aim

Currently there is no gold-standard method for differentiating *S. pseudopneumoniae* from *S. pneumoniae*. However, the correct identification of this organism in clinical isolates is crucial to understand the disease potential of this pathogen and evaluate its significance in diagnostic practice. A simple, unequivocal method to identify *S. pseudopneumoniae* would be valuable. Here we present the use of sequencing the competence gene *comC* as a promising alternative in identifying the pseudopneumococcus.

While determining pherotype diversity in the UK cohort as described in Chapter 5, a rare CSP allele 6.1 was detected. Colonies of this strain, N452, appear small on CBA, was optochin non-susceptible in CO_2 but was susceptible in O_2, bile-insoluble, and was nontypable by anti-capsular antibodies. These phenotypic traits suggest that this strain belongs to *S. pseudopneumoniae*. MLST was performed on this strain and shared 100% identity to *S. pseudopneumoniae* sequences on all seven housekeeping genes. We thus aimed to determine the pherotypes of additional *S. pseudopneumoniae* strains to understand whether CSP6.1 is common in this newly characterized species.

7.2 Materials and Methods

7.2.1 Clinical Specimens and Bacteria

A total of seventeen clinical specimens of *S. pseudopneumoniae*, *S. pneumoniae*, and *S. oralis* were collected at the Royal Free Hospital Microbiology Laboratory between the years 1993-2010 (Table 7.2). Sixteen samples were from patients with lower respiratory tract (LRT) infections, and a single strain was isolated from normally sterile site. Samples were plated on Columbia blood agar (Oxoid, Cambridgeshire) in 5% CO_2 at 35°C in an attempt to cultivate bacteria, and colonies suggestive of pneumococci based on morphology and α-haemolysis were tested with optochin sensitivity. All pseudopneumococci except strain N452 were processed by Dr. Clare Ling. These strains were included in a separate study conducted by Dr. Clare Ling in 2009 prior to the commencement of the work performed by Mr. Marcus Leung in this chapter. Strain N452 was processed by Mr. Marcus Leung. For N452, in addition to α-haemolysis, bile solubility and optochin susceptibility in both ambient and 5% CO_2 conditions were performed.

7.2.2 Genomic DNA Extraction for *comC* Amplification

Genomic DNA from LRT samples were extracted by Dr. Clare Ling using a modified Chelex method as described previously (Walsh, Metzger & Higuchi 1991). The supernatant was used as template for DNA amplification. For culture-positive clinical specimens, genomic DNA was extracted by using the Wizard Genomic DNA Purification Kit (Promega) (by Dr. Clare Ling) or by the heat-lysis method (Section 2.8.1).

7.2.3 Presumptive Identification of *S. pseudopneumoniae*

qPCR for presumptive identification of *S. pseudopneumoniae* was designed and performed by Dr. Clare Ling. Amplifications employing primers and probes specific for Spn9802 (Suzuki *et al.* 2005; Abdeldaim *et al.* 2008) and *lytA* (Carvalho *et al.* 2007) were performed sequentially to differentiate between *S. pneumoniae* (Spn9802 and *lytA*-positive) and *S. pseudopneumoniae* (Spn9802-positive, *lytA*-negative). To assess for PCR inhibition, a potato gene internal amplification control (IAC) was included in each reaction, using primers targeting regulatory photoreceptor gene *phyB* of *Solanum tuberosum* (potato) (Nolan *et al.* 2006).

For *lytA* and Spn9802 qPCR assays, amplification reactions of 25µl were performed containing 1x Platinum® Quantitative PCR SuperMix-UDG (Invitrogen), a final concentration of 7 mM of $MgCl_2$, primers and probes, 4 x 10^{-7} µM IAC template DNA (Sigma-Aldrich), and 5 µl of template DNA (Table 7.3). Negative and positive controls were performed for each qPCR assay. Amplification was performed using a Rotor-Gene Q (Qiagen) with the following conditions: an initial hold cycle at 95 °C for 3 min, followed by 40 cycles of 95 °C for 15 sec and 60 °C for 45 sec. The PCR data was acquired at the end of each cycle and analyzed by the instrument software (Qiagen). Samples with cycle threshold (C_T) value ≤ 35 for the *lytA* target or Spn9802 targets were considered positive; samples with no C_T value for either the *lytA* or Spn9802 targets, but did have a C_T value ≤ 40 for the IAC target were considered negative and samples with a no C_T value for either the *lytA* or Spn9802 targets and no C_T value for the IAD target were considered inhibited. Samples positive for both *lytA* and Spn9802 were considered to be *S. pneumoniae*-positives and samples positive for Spn9802 and negative for *lytA* were considered to be presumptive *S. pseudopneumoniae*-positive. All reactions were performed in triplicate.

Table 7.2. Streptococcal strains included for this study

Species	Strain	Clinical Isolation Site[a]	Accession Number[b]	Source
S. pseudopneumoniae	N452	Blood	Not deposited	This study
	RFH504	LRT	Not deposited	This study
	RFH543	LRT	Not deposited	This study
	RFH686	LRT	Not deposited	This study
	RFH687	LRT	Not deposited	This study
	RFH827	LRT	Not deposited	This study
	RFH905	LRT	Not deposited	This study
	RFH999	LRT	Not deposited	This study
	874	Unknown	AJ240773	(Whatmore, Barcus & Dowson 1999)
	ATCC BAA-960	LRT	Not deposited	This study
	IS7493	LRT	YP004769537	(Shahinas *et al.* 2011)
	PT5479	Nasopharynx	Not deposited	(Simões *et al.* 2010)
	PT5779	Nasopharynx	Not deposited	(Simões *et al.* 2010)
	SK674	Unknown	Not deposited	(Kilian *et al.* 2008)
S. pneumoniae	RFH324	LRT	Not deposited	This study
	RFH410	LRT	Not deposited	This study
	RFH577	LRT	Not deposited	This study
	RFH815	LRT	Not deposited	This study
	RFH864	LRT	Not deposited	This study
	RFH904	LRT	Not deposited	This study
	VA1	Unknown	AJ240789	(Whatmore, Barcus & Dowson 1999)
	41G	Unknown	AJ240766	(Whatmore, Barcus & Dowson 1999)
	CSP2.1b	Nasopharynx	Not deposited	(Vestrheim *et al.* 2011)

	Pn24	Unknown	AJ240759	(Whatmore, Barcus & Dowson 1999)
	Pn59	Unknown	AJ240793	(Whatmore, Barcus & Dowson 1999)
	Pn13	Unknown	AJ240792	(Whatmore, Barcus & Dowson 1999)
	101/87	Unknown	AJ240791	(Whatmore, Barcus & Dowson 1999)
	SK676	Unknown	Not deposited	(Kilian *et al.* 2008)
S. mitis	Col15	Unknown	AJ240762	(Whatmore, Barcus & Dowson 1999)
	Col16	Unknown	AJ240763	(Whatmore, Barcus & Dowson 1999)
	NCTC 10712	LRT	AJ240795	(Whatmore, Barcus & Dowson 1999)
	NCTC 12261	Nasopharynx	AJ000875	(Håvarstein, Hakenbeck & Gaustad 1997)
	B5	Unknown	AJ000871	(Håvarstein, Hakenbeck & Gaustad 1997)
	B6	Unknown	AJ000865	(Håvarstein, Hakenbeck & Gaustad 1997)
	Hu8	Unknown	AJ000866	(Håvarstein, Hakenbeck & Gaustad 1997)
	SK137	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK145	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK262	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK272	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK564	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK596	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK598	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK599	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK601	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK602	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK608	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK609	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK611	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK612	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK614	Unknown	Not deposited	(Kilian *et al.* 2008)

	SK615	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK667	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK675	Unknown	Not deposited	(Kilian *et al.* 2008)
S. oralis	RFH623	LRT	Not deposited	This study
	RFH831	LRT	Not deposited	This study
	Col19	Unknown	AJ240794	(Whatmore, Barcus & Dowson 1999)
	NCTC 11427	Nasopharynx	AJ000873	(Håvarstein, Hakenbeck & Gaustad 1997)
	DSM 20066	Unknown	AJ000874	(Håvarstein, Hakenbeck & Gaustad 1997)
	SK153	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK305	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK34	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK39	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK571	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK597	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK610	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK92	Unknown	Not deposited	(Kilian *et al.* 2008)
S. gordonii	NCTC 3165	Gum	AJ000870	(Håvarstein, Hakenbeck & Gaustad 1997)
	NCTC 7865	Endocardium	X98110	(Håvarstein *et al.* 1996)
	NCTC 7868	Unknown	X98109	(Håvarstein *et al.* 1996)
S. infantis	SK140	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK282	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK283	Unknown	Not deposited	(Kilian *et al.* 2008)
	SK350	Unknown	Not deposited	(Kilian *et al.* 2008)
S. cristatus	NCTC 12479	Unknown	AJ000876	(Håvarstein, Hakenbeck & Gaustad 1997)
S. peroris	ATCC 700780	Tooth	EFX39822	NCBI genome

a LRT: lower respiratory tract

b Accession numbers only available for GenBank submissions over the length of 200 bp. Those sequences under 200 bp are available in respective publications

Table 7.3. Primers and probes used in this study with final reaction concentrations

Targets[a,b]	Forward Primer (Final Concentration)	Reverse Primer (Final Concentration)	Probe (Final Concentration)[c]	Size (bp)	Reference
Spn9802	5'-AGT CGT TCC AAG GTA ACA AGT CT-3' (0.3 µM)	5' ACC AAC TCG ACC ACC TCT TT-3' (0.5 µM)	5' ROX-TAC ATG TAG GAA ACT ATT TTC CTC ACA AA-BHQ2-3' (0.25 µM)	155	(Abdeldaim *et al.* 2008)
lytA	5'- ACG CAA TCT AGC AGA TGA AGC A-3' (0.2 µM)	5'-TCG TGC GTT TTA ATT CCA GCT -3' (0.2 µM)	5' YAK-TGC CGA AAA CGC TTG ATA CAG GGA G-DB-3' (0.2 µM)	75	(Carvalho *et al.* 2007)
IAC	5'-AAC TTG GCT TTA ATG GAC CTC CA-3' (0.25 µM)	5-ACA TTC ATC CTT ACA TGG CAC CA-3 (0.25 µM)	5' Cy5-TGC ACA AGC TAT GGA ACA CCA CGT- BBQ-3' (200 nM)	101	(Nolan *et al.* 2006)
comC[†]	5'-GAT AAA ATT CTC TCA ACT GT-3' (0.4 µM)	5'-GGT AAC TGT GAC TAA TAA TT-3' (0.4 µM)	None	~300	This study

[a] IAC: internal amplification control based on method described previously (Nolan, 2006). IAC template amplicon sequence as follows: 5'-AAC TTG GCT TTA ATG GAC CTC CAA TTT TGA GTG TGC ACA AGC TAT GGA ACA CCA CGT AAG ACA TAA AAC GGC CAC ATA TGG TGC CAT GTA AGG ATG AAT GT-3'

b *comC* was targeted using conventional PCR without a probe

c BHQ: Black Hole Quencher; YAK: Yakima Yellow; DB: Dabcyl; BBQ: Black Berry Quencher

7.2.4 Amplification and Sequence Analysis of *comC*

PCR amplification, purification, cycle sequencing, and sequence analysis were performed as described in Sections 2.8.2 to 2.8.8. Primers specific for *comC* were employed for both PCR and cycle sequencing. Please refer to Table 2.6 for primer sequences and annealing temperature.

7.2.5 Multilocus Sequence Typing (MLST)

The primers for MLST used are as described in Table 2.6, with the same reaction components and concentrations as indicated in Chapter 4. Sequences were uploaded onto GenBank database with accession numbers JQ352770-JQ352776.

7.2.6 Publicly Accessible Sequences of Streptococcal CSP

Additional streptococcal ComC amino acid sequences were included in this study for the construction of a phylogenetic tree (see below). Sequences obtained from previous publications with accession numbers are indicated in Table 7.2. Sequences with no accession numbers are provided in respective publications.

7.2.7 Construction of a ComC Phylogenetic Tree

Multiple alignment of *comC* sequences were constructed and translated to amino acid sequences on MEGA version 5.05 (Tamura *et al.* 2011). For both analyses neighbour-joining trees were constructed based on amino acid alignment data. Bootstrap support of 1,000 repetitions with replacements was performed.

7.3 Results

7.3.1 *S. pseudopneumoniae* ComC Sequences

It was possible to identify ComC sequences for nine pseudopneumococcal strains where CSP sequences were derived directly from samples submitted to our laboratory with LRT and invasive infections. These samples were presumed to contain pseudopneumococci based on a Spn9802-positive and *lytA*-negative qPCR reaction (eight strains), bile-insolubility and optochin-variable phenotypic traits (one strain, N452), and/or MLST (one strain, N452). The pherotype of control strain ATCC BAA-960 was also characterized in this study. Two pherotypes were detected in these nine strains, six (N452, RFH504, RFH543, RFH687, RFH905, RFH999) of which were associated with CSP6.1. To avoid confusion with pherotypes of other oral streptococci we propose that CSP6.1 be named as CSPps1, where "ps" represents the pseudopneumococcus. The three remaining strains, BAA-960, RFH686 RFH827, had a ComC sequence that has not been reported before which is identical to CSPps1 in size, differing by an alanine → serine substitution at position 12 of the propeptide. We propose that this pherotype be classified as CSPps2 (Fig. 7.2). Nucleotide sequences of *comC* for strains characterized in this study are indicated in Table 7.4. All strains of a pherotype had the identical nucleotide sequence.

```
CSP6.1      MK-NT--VKLEQFVALKEKDLQKIKGGEMRLPKILRDFIFPRKK 41
CSP6.3      MK-NT--VKLEQFVSLKEKDLQKIKGGEMRLPKILRDFIFPRKK 41
-----------------------------------------------------
SK674       MKKNTDFAQMKDFQQLNEKELQEIRGGEWRPPYTINNFLFPKRK 44
SK350       MKKHTGFAQMKDFQELNEKELQEIRGGEWRPMYTINNFLFSKSK 44
```

Fig. 7.2. Pherotypes of *S. pseudopneumoniae* CSP6.1 (CSPps1), CSP6.3 (CSPps2), and SK674. Mature region of peptide is in bold after the double-glycine. CSP6.1 and CSP6.3 differ by a single amino acid at position 12 (alanine in CSP6.1 and serine in CSP6.3), as shaded in yellow. SK674 is a presumptive pseudopneumococcus (Kilian *et al.* 2008), with an extended ComC compared to CSP6.1 and CSP6.3, and shares higher identity to *S. infantis* SK350. A total of three pseudopneumococcal pherotypes were characterized. Differences between SK674 and CSP6.1/6.3 present in the mature peptide region are shaded in light blue. Double-glycine motif shaded in red.

Table 7.4 Nucleotide sequences of pherotypes characterized in this study.

Pherotype	Nucleotide sequence (5'-3')
CSP6.1 (CSPps1)	atgaaaaaca cagttaaatt ggaacagttt gtagccttga aggaaaaaga cttgcaaaag attaaaggtg gggaaatgag acttccaaaa atcctccgtg attttatttt cccaagaaaa aagtaa
CSP6.3 (CSPps2)	atgaaaaaca cagttaaatt agaacagttt gtatctttga aggaaaaaga cttgcaaaag gggaaatgag acttccaaaa atcctccgtg attttatttt cccaagaaaa aagtaa
CSP1	atgaaaaaca cagttaaatt ggaacagttt gtagctttga aggaaaaaga cttacaaaag attaaaggtg gggagatgag gttgtcaaaa ttcttccgtg attttatttt acaaagaaaa aagtaa
CSP6.2	atgaaaaaca cagttaaatt ggaacagttt gtagccttga aggaaaaaga cttgcaggag attaaaggtg gggagatgag actgccaaaa atcctccgtg attttatttt cccaagaaaa aagtaa

CSP sequences of four strains (IS7493, PT5479, PT5779, and SK674) were available from reports in earlier publications (Table 7.2). All but SK674 were identical to CSPps1. Thus, more than 70% of the presumptive *S. pseudopneumoniae* strains in this study were associated with CSPps1 (Table 7.5). SK674 has an extended ComC at 54 amino acids, most similar to a pherotype characterized in *Streptococcus infantis* SK350, with six amino acid substitutions, three of which in the mature peptide region (Fig. 7.2). SK674 differ from CSPps1 and CSPps2 in most of the mature peptide region, including positions 4, 7, 8, and 11, which had been hypothesized to confer competence receptor specificity (Johnsborg *et al.* 2006).

Table 7.5. Distribution of Pseudopneumococcal Pherotypes

Pherotype/ Strain	**Proportion (% of n = 14)**	**Strains**
CSPps1 (CSP6.1)	10 (71)	N452, RFH504, RFH543, RFH687, RFH905, RFH999, IS7493, 874, PT5479, PT5779
CSPps2 (CSP6.3)	3 (21)	BAA-960, RFH686, RFH827
SK674	1 (7)	SK674

7.3.2 Phylogenetic Analysis of Streptococcal Pherotypes

A phylogenetic tree constructed from alignment of streptococcal pherotypes shows that, by pherotype, streptococci fall into two major groups (Fig. 7.3). *S. pneumoniae*, *S. pseudopneumoniae*, *S. mitis* and some *S. oralis* strains belonged to one group (group 1), while a more divergent and loosely-defined group (group 2) consists of predominantly *S. gordonii*, *S. infantis*, *S. peroris*, *S. cristatus*, and most of the remaining *S. oralis* strains. SK674, formerly classified as a pseudopneumococcus (Kilian *et al.* 2008), clustered near members of *S. infantis* by ComC alignment. All of the remaining *S. pseudopneumoniae* pherotypes are grouped in a separate cluster in close relation to other species, notably *S. oralis* and *S. mitis* (Fig. 7.3).

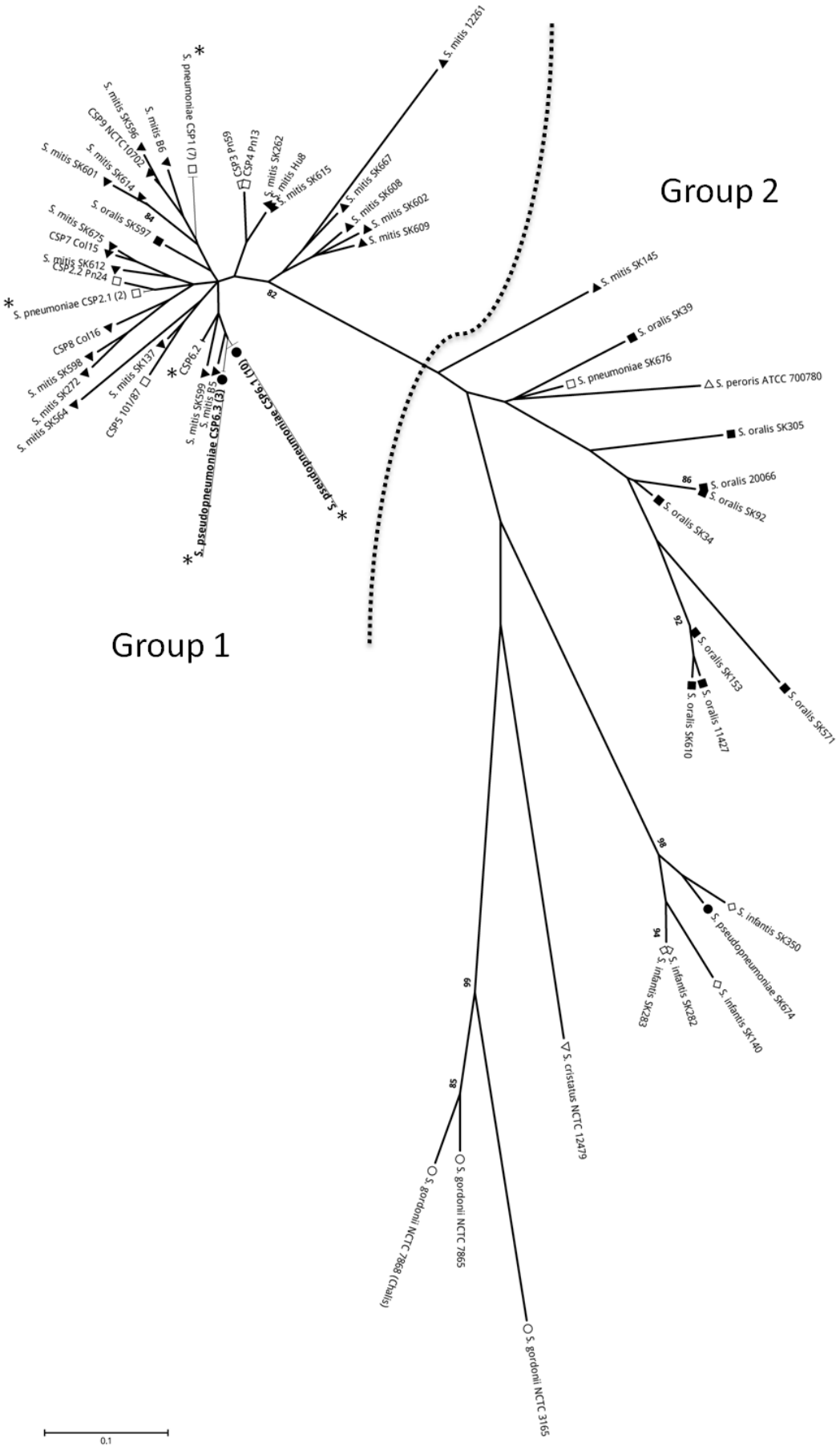

Group 2
Group 1
S. mitis 12261
S. pneumoniae CSP1 (7)
S. mitis SK596
S. mitis B6
CSP9 NCTC10702
S. mitis SK614
S. mitis SK601
CSP3 Pn59
CSP4 Pn13
S. mitis SK262
S. mitis Hu8
S. mitis SK615
S. mitis SK667
S. mitis SK608
S. mitis SK602
S. mitis SK609
84
S. oralis SK597
S. mitis SK675
CSP7 Col15
S. mitis SK612
CSP2.2 Pn24
S. pneumoniae CSP2.1 (2)
82
CSP8 Col16
S. mitis SK598
S. mitis SK272
S. mitis SK564
S. mitis SK137
CSP5 101/87
CSP6.2
S. mitis SK599
S. mitis B5
S. pseudopneumoniae CSP6.3 (3)
S. pseudopneumoniae CSP6.1 (10)
S. mitis SK145
S. oralis SK39
S. pneumoniae SK676
S. peroris ATCC 700780
S. oralis SK305
86
S. oralis 20066
S. oralis SK92
S. oralis SK34
92
S. oralis SK153
S. oralis SK610
S. oralis 11427
S. oralis SK571
98
S. infantis SK350
S. pseudopneumoniae SK674
94
S. infantis SK283
S. infantis SK282
S. infantis SK1140
66
S. cristatus NCTC 12479
85
S. gordonii NCTC 7868 (Challis)
S. gordonii NCTC 7865
S. gordonii NCTC 3165
0.1

Fig. 7.3. Neighbour-joining phylogenetic representation of Streptococcal ComC. Two clusters of ComC based on amino acid sequence similarities (groups 1 and 2) are separated by a curved dotted line. Each pherotype is labelled according to the classified species: filled circle, *S. pseudopneumoniae*; open square, *S. pneumoniae*; filled upward triangle, *S. mitis*; filled square, *S. oralis*; open circle, *S. gordonii*; open upward triangle, *S. peroris*; open diamond, *S. infantis*, open downward triangle, *S. cristatus.* Groups with asterisk indicate multiple strains of the same pherotype that have been compressed for clarity. Numbers of strains with same pherotype are indicated in parentheses. Pseudopneumococcal strains of these groups are indicated in Table 2. Pneumococcal CSP1 pherotype includes strains VA1, RFH324, RFH410, RFH577, RFH815, RFH864, and RFH904. Pneumococcal CSP2.1 pherotype includes strains 41G and CSP2.1b. CSP6.2 pherotypes include SK671 (*S. mitis*), Col19 (*S. oralis*), RFH623 (*S. oralis*), and RFH831 (*S. oralis*). Phylogenetic tree was built with 1,000 bootstrap repetitions with replacement, with support over 80% indicated. Ruler indicates amino acid substitutions per site. Tree constructed with MEGA 5.05

7.4 Discussion

7.4.1 CSP6.1 is a Common Pseudopneumococcal Pherotype

We have characterized pherotypes associated with *S. pseudopneumoniae* by comparing strains available to us, amplifying sequence collected from lower respiratory tract and blood samples and collecting publicly available ComC sequences for this organism and related streptococcal species. We have shown that CSP6.1 (or CSPps1) is the commonest pherotype among *S. pseudopneumoniae* found in different geographical regions. Pherotype CSP6.1 was previously considered to be a rare pherotype of an "atypical nontypable pneumococcal" strain 874 based on multilocus sequence analysis (Müller-Graf *et al.* 1999; Whatmore, Barcus & Dowson 1999). Its classification as a pneumococcus may stem from its possession of *ply* and *lytA*, which were once considered suitable genetic markers for the pneumococcus (Müller-Graf *et al.* 1999). However, it is known that these two genes are not specific to *S. pneumoniae* (Whatmore *et al.* 2000; Neeleman *et al.* 2004; Johnston *et al.* 2010), and we have been unable to find a report of CSP6.1 being found in a strain unequivocally identified as *S. pneumoniae*. Based on these observations, we hypothesize that strain 874 is a member of *S. pseudopneumoniae.* Strain 101/87 associated with CSP5 was described as an "atypical pneumococcus" and could not be serotyped by Whatmore *et al.* (Whatmore, Barcus & Dowson 1999). Phylogenetic analysis of CSP sequences in this study suggests that CSP5 is most closely related to *S. mitis* or *S. oralis* pherotypes that cluster together. Thus, we believe that CSP 6.1 is associated with *S. pseudopneumoniae* and to differentiate pseudopneumococcal pherotypes from ComC sequences from other organisms, they be designated CSPps1 instead of CSP6.1. In this study we have identified a new pherotype associated with *S. pseudopneumoniae* and have designated these as CSPps2. *S. pseudopneumoniae* pherotypes form a distinct cluster within those of other oral streptococcal species, suggesting that pseudopneumococcal pherotypes could be species-specific and be used as a simple diagnostic tool.

7.4.2 Strain SK674 with a Divergent ComC

In contrast, strain SK674, a *S. pseudopneumoniae* strain identified based on clustering of housekeeping gene sequences (Kilian *et al.* 2008), clustered closely in ComC sequences of *S. infantis* strains. One might argue that SK674 acquired a divergent *comC* from *S. infantis* by horizontal gene transfer. While interspecies transfer of the

competence operon has been documented (Håvarstein, Hakenbeck & Gaustad 1997), the lack of such an event in a collection of over 200 pneumococcal strains indicates that this is relatively rare. SK674 has an estimated genome size of 1.87 Mbp (Kilian *et al.* 2008), comparable to that of other *S. infantis* strains (1.74-1.88 Mbp), and much smaller than the genome size of the pseudopneumococcal strain IS7493 (2.1 Mbp) (Shahinas *et al.* 2011). It seems that the alternative explanation, that SK674 is actually a strain of *S. infantis* is more likely. Additional analysis of SK674, such as DNA-DNA hybridization and *lytA* sequence analysis, may shed light as to the true identity of this strain.

7.4.3 Competence Gene *comC* in Genome of IS7493

The recently characterized genome sequence of *S. pseudopneumoniae* strain IS7493 concluded that CSP-mediated induction of fratricide does not take place based on the absence of *comC* in this strain (Shahinas *et al.* 2011). We were, however able to locate the gene that is identical to CSPps1. Thus, we can conclude that this *S. pseudopneumoniae* strain does contain the necessary gene sequences for production of a competence peptide, and that a homologous gene is found in all pseudopneumococcal strains analyzed thus far, strengthening the use of this gene as a diagnostic marker.

7.4.4 ComC as a Potential Target for Pseudopneumococcal Identification

S. pseudopneumoniae is usually identified as acapsulate, bile-insoluble, and optochin intermediate-resistant in 5% CO_2 and optochin susceptible in ambient O_2 (Arbique *et al.* 2004), tests that can be difficult to standardize in the laboratory. Identification by *comC* sequencing would allow a rapid method of definitive diagnosis as competence ligand gene is conserved across streptococcal species, pseudopneumococcal *comC* sequences appear to provide taxonomic information. One of the limitations of this study is the low number of strains, and follow-up studies employing a larger collection of *S. pseudopneumoniae* strains would provide greater understanding to the potential of this gene as a species identification target.

A recent report suggested that sequencing *recA* could differentiate between *S. pneumoniae* from *S. pseudopneumoniae* (Sistek *et al.* 2011), but the study was performed with a smaller number of pseudopneumococcal strains solely from North

America. The detection of homologues of pneumococcal virulence factors in *S. pseudopneumoniae* and the recent detection of pseudopneumococci in carriage with antibiotic resistance necessitate its accurate diagnosis as an emerging causative agent of diseases (Keith & Murdoch 2008; Johnston *et al.* 2010). Here we propose that pherotyping may be a promising diagnostic alternative based on the clustering of pseudopneumococcal pherotypes from different continents.

7.4.5 ComC and Streptococcal Species Diversity

As most groups of strains within a species communicate in a concerted manner by induction of competence by the same CSP allele, pherotype diversification may thus be one of the initial steps for speciation. Indeed, studies have indicated that strains of *S. mitis* are no more related to each other than between strains of *S. mitis* and *S. pneumoniae* based on MLSA, and this is reflected on the great pherotype diversity of the former species (Kilian *et al.* 2008). Similarly, it may be that *S. pseudopneumoniae* acquired, either through genetic drift or horizontal gene transfer, a variant pherotype, and had formed an exclusive group in which communication (in this case, competence induction) occurred within this small group. This subsequently led to diversification of genetic traits elsewhere in the genome. While competence and genetic exchange in this emerging pathogen has not been demonstrated *in vivo*, this divergent species may also expand the streptococcal pan-genome by two mechanisms: pseudopneumococcal intra-species diversification and horizontal gene exchange between members of the pseudopneumococcus and related members of the Mitis streptococcal group.

The establishment of *S. pseudopneumoniae* as a potential donor and/or recipient in inter-species gene exchange is crucial to the greater understanding of this organism's role in the streptococcal supragenome. Co-culture experiments between *S. pseudopneumoniae* and *S. pneumoniae* and *S. mitis* with selection marker would indicate whether *S. pseudopneumoniae* participates in genetic exchanges with its close bacterial relatives.

7.5 Concluding Remarks

In conclusion, we propose CSP sequence analysis can provide rapid accurate differentiation of *S. pseudopneumoniae* from closely related species *S. pneumoniae, S. mitis,* and *S. oralis*. With this in mind, we anticipate that some strains currently classified as atypical pneumococci can be identified as pseudopneumococci based on ComC sequencing. Use of ComC sequencing will simplify the gathering of data to understand the disease potential of this organism that may be now emerging as a pathogen. To add to this observation, we would encourage other laboratories to sequence "atypical pneumococcal" strains to provide the more sequences to confirm whether *comC* may be used as a rapid marker for the identification of this emerging pathogen.

CHAPTER EIGHT: Final Discussion and Future Work

Works included in this thesis included the design of a novel single-primer PCR and sequencing-based sequetyping method to correctly identify serotypes of commonly encountered *S. pneumoniae* strains. Through this we have identified extensive sequence variations at the regulatory gene *cpsB* in both species and serotype level. We have employed this method to assist in understanding the adaptive potential of the pneumococcal co-colonization in healthy children in Tanzania. Previous works analyzing the pneumococcal supragenome have revealed extensive heterogeneity in gene content amongst pneumococci, even within a single ST (Shen *et al.* 2006; Silva *et al.* 2006; Hiller *et al.* 2007; Dagerhamn *et al.* 2008; Donati *et al.* 2010; Hiller *et al.* 2010; Croucher *et al.* 2011; Hiller *et al.* 2011). Here we have provided a framework in understanding the supragenome size within the nasopharynx. We observed that most pneumococci in co-colonization events are of different competence pherotypes, thus providing evidence that the paediatric upper respiratory tract is an optimum environment for pneumococcal gene transfer. We also suggest that CSP6.1, which was detected previously as a pherotype of an "atypical pneumococcus," is a common CSP allele of *S. pseudopneumoniae*.

8.1 Sequetyping of Pneumococci

The study described the design and application of a novel, single primer pair-based serotype identification method. Based on *in silico* analysis, PCR targeting the capsulation regulatory gene *cpsB* enabled successful species-specific amplification of a large number of pneumococcal serotypes; all serotypes in the latest 13-valent conjugate vaccines and common replacement serotypes could be identified accurately at least to the serogroup level. A recent global assessment of prevalent serotypes indicates that the novel method should correctly identify at least 70% of pneumococci causing invasive diseases in children under the age of 5 worldwide (Johnson *et al.* 2010). Recent epidemiological accounts of invasive serotypes indicate that virtually all of invasive pneumococci in Tanzanian children are PCV13 serotypes (Mudhune, Wamae & Region 2009; Crump *et al.* 2011a). Sequetyping is also more economical than conventional Quellung serotyping. However, the efficacy of sequetyping has not been assessed in all known pneumococcal serotypes; the results presented included

only 48 serotypes. In addition, a large number of serotypes contained only a single strain, and reproducibility of sequetyping for these serotypes could not be verified. There is also the difficulty in detecting multiple colonization events, as sequence data would not distinguish respective sequences.

8.1.1 Future Work

Our research group is currently acquiring additional strains of the remaining serotypes not included, and also those with only one strain to access reproducibility of these serotypes. Inclusion of additional strains would also address serotypes with discrepant *in silico* and *in vitro* results, as well as misidentified strains in this study. A blind-study is currently undergoing to assess accuracy of sequetyping from polymicrobial samples.

8.2 Co-Colonization of *S. pneumoniae* Strains in Tanzanian Children

In Chapter 4 we described the multiple colonization in Tanzanian children under the age of 6, illustrating the adaptive potential of this organism in the nasopharynx where horizontal gene transfer is thought to take place. Characterization of up to twenty pneumococcal colonies from nasopharyngeal and orophayngeal swabs of children revealed the simultaneous colonization of strains of different serotypes, antibiotic susceptibilities, and genetic backgrounds. We also provided direct evidence that assessing serotype diversity alone is an underestimate of strain diversity, as pneumococci expressing the same capsule may have different STs in the upper respiratory tract.

S. pneumoniae is among the most common detectable etiological agents of bloodstream (Matee & Matre 2001; Crump *et al.* 2011a; Crump *et al.* 2011b; Thriemer *et al.* 2012) and respiratory infections (Uriyo *et al.* 2006) in Tanzania, and high proportions of non-susceptibility to penicillin and co-trimoxazole have been reported in this country (Batt *et al.* 2003; Moyo *et al.* 2012). The co-colonization of strains with different susceptibilities to both drugs is likely to facilitate the spread of nonsusceptible pneumococci.

However, the supragenome size within a single host at one time is yet to be determined. This is likely to be dependent on the cohort studied and the strain

diversity of the nasopharynx, as the supragenome size is likely to increase with additional strains. In our study, child 35 was presented with total of six strains co-colonizing, which according to Hiller's finite supragenome model increases the gene repertoire for HGT within the nasopharynx by 30% (Hiller *et al.* 2007).

8.2.1 Future Work

Whole-genome sequencing of strains within child 35, who showed the highest number (six) of multiple strains co-colonizing, would provide an approximation of the supragenome size available for HGT within a host at one time. The number of core genes present in all strains, distributed genes in some strains, and unique genes in single strains can be determined, and compared to previous works where supragenome as a species in general was assessed.

Due to the limitation of time and resources, MLST was only performed in colonization events of multiple phenotypes and of selected single colonization events. Although we were unable to detect presence of different strains within single colonization events, it is possible that children had been colonized by different strains that were identical in serotype and antibiotic susceptibilities. Therefore, the described rate of co-colonization is likely to be an underestimate of the cohort. Determining the ST of strains in other colonization episodes would provide a more in-depth analysis of such colonization events.

Furthermore, comparison of supragenome sizes of single and multiple colonizations is of interest. Dagerhamn *et al.* (Dagerhamn *et al.* 2008) showed that strains of the same or related ST are more likely to share a higher number of genes in their genomes. Understanding the supragenome sizes of multiple and single colonizations would therefore directly indicate how multiple colonization attributes to pneumococcal adaptation.

8.3 Diversity of CSP Alleles in the UK and Tanzania

We have revealed differences in proportions of pherotype distributions between two countries (UK and Tanzania). This is the first report of pherotype diversity in a developing nation, and the first to report of a high prevalence of CSP4. It is likely that pherotype distributions are governed by geographical differences and STs, as this and

other studies have demonstrated the close association between pherotype and clonal complexes (Carrolo *et al.* 2009; Vestrheim *et al.* 2011). In addition, co-colonization events commonly contain strains of different pherotypes, thereby potentially facilitating genetic exchanges through competence-mediated fratricide.

8.3.1 Future Work

The disease potential of CSP4 is not clear. In this study, CSP4 pneumococci were detected in both the nasopharynx and blood. Animal models involving strains of CSP4 would allow a greater understanding of host-pathogen interaction in this CSP allele. It is of interest to point out that Iannelli detected *comC1*-expressing strains with a *comD4* receptor, and that this receptor can activate competence at elevated concentrations (~30 ng/mL) of CSP1.

8.4 Intra-Serotype Diversity of *cpsB*

Comparison of multiple strains of single serotypes revealed that there exist significant differences in intra-serotype diversity in *cpsB* between serotypes. One of the limitations of this work is that MLST data was not available for all strains, and comparison between more serotypes could not be performed. However, MLST data for a number of serotype 6B and serotype 1 strains was available, and we observed the differences in intra-serotype diversity of *cpsB* that is independent of genetic relatedness as determined by MLST.

8.4.1 Future Work

The inclusion of additional serotypes and multiple strains of studied serotypes, which is addressed in Section 8.1, would allow the analysis of intra-serotype *cpsB* diversity to be compared between additional serotypes.

Two clusters of divergent serotype 6B strains based on *cpsB* sequences were observed. However, it is important to understand whether the clusters account for phenotypic differences, such as the amount of capsule produced and the rate of enzymatic activity. In order to perform such a comparison, characterization of the remaining genes present in the *cps* locus is required for these strains, as sequence variation may be found elsewhere along *cps* (Morona, Morona & Paton 1999a; Varvio *et al.* 2009; Elberse *et al.* 2011). Comparison of capsule production and invasive

potential (with the use of animal models) of isogenic mutants differing in the *cps* region between clusters within a given serotype would provide a biological relevance to the observed sequence variations.

Amino acid analysis revealed the presence of non-conservative substitution hotspots within α4 and α7 globular domains (Hagelueken *et al.* 2009). However, it is not known whether structural changes would entail as a result of these substitutions. Structural modeling and phosphatase activity assays between strains of blue and red serotypes, or of different sequence clusters within serotype 6B, would provide a relationship between between sequence variation and structural differences and biological activity.

8.5 ComC as a Target for Identifying the Pseudopneumococcus

The discovery of an "atypical pneumococcus" associated with CSP6.1 led to its phenotypic characterization, and this strain was subsequently identified as *S. pseudopneumoniae*. Analysis of additional pseudopneumococcal strains based on presence and absence of Spn9802 and *lytA*, respectively, provided evidence that CSP6.1 is a common pseudopneumococcal pherotype as supposed to a rare pneumococcal pherotype as suggested (Whatmore, Barcus & Dowson 1999). A greater understanding the diversity of both pseudopneumococcal pherotypes and that of the *Streptococcus* would

8.5.1 Future Work

The only other CSP6.1 strain detected previously was a nontypable strain (strain 874) from Kenya, and intriguingly strain 874 had identical sequences in *comA* to other pneumococci. Therefore, pherotype identification of additional pseudopneumococcal strains is encouraged, and analysis of other competence genes would provide additional information on the competence machinery of this organism. In addition, mixed culture experiments involving pseudopneumococci of different pherotypes, as well as different streptococcal species may reveal the role of *S. pseudopneumoniae* in inter-specific genetic exchange and also streptococcal species diversification.

References

Aanensen, D. M., A. Mavroidi, S. D. Bentley, P. R. Reeves & B. G. Spratt. 2007, "Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci", *J. Bacteriol.*, vol. 189, no. 21, 7856-7876.

Abdeldaim, G. M. K., K. Strålin, P. Olcén, J. Blomberg & B. Herrmann. 2008, "Toward a quantitative DNA-based definition of pneumococcal pneumonia: a comparison of *Streptococcus pneumoniae* target genes, with special reference to the Spn9802 fragment", *Diagn. Microbiol. Infect. Dis.*, vol. 60, no. 2, 143-150.

Abdullahi, O., J. Nyiro, P. Lewa, M. Slack & J. Scott. 2008, "The descriptive epidemiology of *Streptococcus pneumoniae* and *Haemophilus influenzae* nasopharyngeal carriage in children and adults in Kilifi district, Kenya", *Pediatr. Infect. Dis. J.*, vol. 27, no. 1, 59.

Abdullahi, O., E. Wanjiru, R. Musyimi, N. Glass & J. A. G. Scott. 2007, "Validation of nasopharyngeal sampling and culture techniques for detection of *Streptococcus pneumoniae* in children in Kenya", *J. Clin. Microbiol.*, vol. 45, no. 10, 3408-3410.

Adrian, P. & K. Klugman. 1997, "Mutations in the dihydrofolate reductase gene of trimethoprim-resistant isolates of *Streptococcus pneumoniae*", *Antimicrob. Agents Chemother.*, vol. 41, no. 11, 2406.

Alcantara, R., L. Preheim & M. Gentry-Nielsen. 2001, "Pneumolysin-induced complement depletion during experimental pneumococcal bacteremia", *Infect. Immun.*, vol. 69, no. 6, 3569.

Alexander, J. E., R. A. Lock, C. C. Peeters, J. T. Poolman, P. W. Andrew, T. J. Mitchell, D. Hansman & J. C. Paton. 1994, "Immunization of mice with pneumolysin toxoid confers a significant degree of protection against at least nine serotypes of *Streptococcus pneumoniae*.", *Infect. Immun.*, vol. 62, no. 12, 5683-5688.

Allan, E., H. A. Hussain, K. R. Crawford, S. Miah, Z. K. Ascott, M. H. Khwaja & A. H. Hosie. 2007, "Genetic variation in *comC*, the gene encoding competence-stimulating peptide (CSP) in *Streptococcus mutans*", *FEMS Microbiol. Lett.*, vol. 268, no. 1, 47-51.

Allegrucci, M., F. Hu, K. Shen, J. Hayes, G. Ehrlich, J. Post & K. Sauer. 2006, "Phenotypic characterization of *Streptococcus pneumoniae* biofilm development", *J. Bacteriol.,* vol. 188, no. 7, 2325-2335.

Allegrucci, M. & K. Sauer. 2007, "Characterization of colony morphology variants isolated from *Streptococcus pneumoniae* biofilms.", *J. Bacteriol.,* vol. 189, no. 5, 2030-2038.

Allegrucci, M. & K. Sauer. 2008, "Formation of *Streptococcus pneumoniae* non-phase-variable colony variants is due to increased mutation frequency present under biofilm growth conditions", *J. Bacteriol.,* vol. 190, no. 19, 6330-6339.

Alloing, G., B. Martin, C. Granadel & J. Claverys. 1998, "Development of competence in *Streptococcus pneumoniae*: pheromone autoinduction and control of quorum sensing by the oligopeptide permease", *Mol. Microbiol.,* vol. 29, no. 1, 75-83.

Andrade, A. L. S., C. M. Franco, J. Lamaro-Cardoso, M. C. D. André, L. L. Oliveira, A. Kipnis, C. G. Rocha, J. G. Andrade, S. L. Alves, I. H. Park, M. H. Nahm, S. G. Almeida & M. C. C. Brandileone. 2010, "Non-typeable *Streptococcus pneumoniae* carriage isolates genetically similar to invasive and carriage isolates expressing capsular type 14 in Brazilian infants", *J. Infect.*, vol. no. 1-9.

Antonio, M., I. Hakeem, K. Sankareh, Y. B. Cheung & R. A. Adegbola. 2009, "Evaluation of sequential multiplex PCR for direct detection of multiple serotypes of *Streptococcus pneumoniae* from nasopharyngeal secretions", *J. Med. Microbiol.*, vol. 58, no. 3, 296-302.

Appelbaum, P., J. Scragg, A. Bowen, A. Bhamjee, A. Hallett & R. Cooper. 1977, "*Streptococcus pneumoniae* resistant to penicillin and chloramphenicol", *Lancet,* vol. 310, no. 8046, 995-997.

Aravind, L. & E. Koonin. 1998, "Phosphoesterase domains associated with DNA polymerases of diverse origins", *Nucleic Acids Res.,* vol. 26, no. 16, 3746-3752.

Arbique, J., C. Poyart, P. Trieu-Cuot, G. Quesne, M. Carvalho, A. Steigerwalt, R. Morey, D. Jackson, R. Davidson & R. Facklam. 2004, "Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae* sp. nov.", *J. Clin. Microbiol.*, vol. 42, no. 10, 4686-4696.

Arifeen, S. E., S. K. Saha, S. Rahman, K. M. Rahman, S. M. Rahman, S. Bari, A. Naheed, I. Mannan, M. H. R. Seraji, N. U. Ahmed, M. S. Hassan, N. Huda, A. U. Siddik, I. Quasem, M. Islam, K. Fatima, H. Al Emran, W. A. Brooks, A. H. Baqui, R. F. Breiman, D. Sack & S. P. Luby. 2009, "Invasive pneumococcal disease among children in rural Bangladesh: results from a population-based surveillance.", *Clin. Infect. Dis.,* vol. 48 Suppl 2, no. S103-113.

Arrecubieta, C., E. Garcia & R. López. 1995, "Sequence and transcriptional analysis of a DNA region involved in the production of capsular polysaccharide in *Streptococcus pneumoniae* type 3", *Gene,* vol. 167, no. 1-2, 1-7.

Arrecubieta, C., R. López & E. García. 1994, "Molecular characterization of *cap3A*, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae* type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome", *J. Bacteriol.,* vol. 176, no. 20, 6375-6383.

Auranen, K., J. Mehtälä, A. Tanskanen & M. S. Kaltoft. 2010, "Between-strain competition in acquisition and clearance of pneumococcal carriage--epidemiologic evidence from a longitudinal study of day-care children", *Am. J. Epidemiol.,* vol. 171, no. 2, 169-176.

Austrian, R. 1976, "The quellung reaction, a neglected microbiologic technique.", *Mt. Sinai J. Med.,* vol. 43, no. 6, 699-709.

Austrian, R. 1981, "Pneumococcus: the first one hundred years", *Rev. Infect. Dis.,* vol. 3, no. 2, 183-189.

Avery, O. T., C. M. Macleod & M. McCarty. 1944, "Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III", *J. Exp. Med.,* vol. 79, no. 2, 137-158.

Azzari, C., M. Moriondo, G. Indolfi, M. Cortimiglia, C. Canessa, L. Becciolini, F. Lippi, M. De Martino & M. Resti. 2010, "Realtime PCR is more sensitin pneumococcal invasive disease", *PLoS ONE,* vol. 5, no. 2, e9282.

Azzari, C., M. Moriondo, G. Indolfi, C. Massai, L. Becciolini, M. De Martino & M. Resti. 2008, "Molecular detection methods and serotyping performed directly on clinical samples improve diagnostic sensitivity and reveal increased incidence of invasive disease by *Streptococcus pneumoniae* in Italian children", *J. Med. Microbiol.,* vol. 57, no. 10, 1205-1212.

Bagnoli, F., M. Moschioni, C. Donati, V. Dimitrovska, I. Ferlenghi, C. Facciotti, A. Muzzi, F. Giusti, C. Emolo, A. Sinisi, M. Hilleringmann, W. Pansegrau, S. Censini, R. Rappuoli, A. Covacci, V. Masignani & M. A. Barocchi. 2008, "A second pilus type in *Streptococcus pneumoniae* is prevalent in emerging serotypes and mediates adhesion to host cells", *J. Bacteriol.*, vol. 190, no. 15, 5480-5492.

Balachandran, P., S. K. Hollingshead, J. C. Paton & D. E. Briles. 2001, "The autolytic enzyme LytA of *Streptococcus pneumoniae* is not responsible for releasing pneumolysin", *J. Bacteriol.*, vol. 183, no. 10, 3108-3116.

Balakrishnan, I., P. Crook, R. Morris & S. H. Gillespie. 2000, "Early predictors of mortality in pneumococcal bacteraemia", *J. Infect.*, vol. 40, no. 3, 256-261.

Balsalobre, L., M. J. Ferrándiz, J. Liñares, F. Tubau & A. G. de la Campa. 2003, "Viridans group streptococci are donors in horizontal transfer of topoisomerase IV genes to *Streptococcus pneumoniae*", *Antimicrob. Agents Chemother.*, vol. 47, no. 7, 2072-2081.

Balsalobre, L., A. Hernández-Madrid, D. Llull, A. J. Martín-Galiano, E. García, A. Fenoll & A. G. de la Campa. 2006, "Molecular characterization of disease-associated streptococci of the mitis group that are optochin susceptible.", *J. Clin. Microbiol.*, vol. 44, no. 11, 4163-4171.

Barnes, D., S. Whittier & P. Gilligan. 1995, "Transmission of multidrug-resistant serotype 23F *Streptococcus pneumoniae* in group day care: evidence suggesting capsular transformation of the resistant strain *in vivo*", *J. Infect.*, vol. 171, no. 4, 890-896.

Barocchi, M. A., J. Ries, X. Zogaj, C. Hemsley, B. Albiger, A. Kanth, S. Dahlberg, J. Fernebro, M. Moschioni, V. Masignani, K. Hultenby, A. R. Taddei, K. Beiter, F. Wartha, A. von Euler, A. Covacci, D. W. Holden, S. Normark, R. Rappuoli & B. Henriques-Normark. 2006, "A pneumococcal pilus influences virulence and host inflammatory responses", *Proc. Nat. Acad. Sci. U. S. A*, vol. 103, no. 8, 2857-2862.

Bartilson, M., A. Marra, J. Christine, J. S. Asundi, W. P. Schneider & A. E. Hromockyj. 2001, "Differential fluorescence induction reveals *Streptococcus pneumoniae* loci regulated by competence stimulatory peptide", *Mol. Microbiol.*, vol. 39, no. 1, 126-135.

Bast, D. J., D. E. Low, C. L. Duncan, L. Kilburn, L. A. Mandell, R. J. Davidson & J. C. de Azavedo. 2000, "Fluoroquinolone resistance in clinical isolates of *Streptococcus pneumoniae*: contributions of type II topoisomerase mutations and efflux to levels of resistance", *Antimicrob. Agents Chemother.*, vol. 44, no. 11, 3049-3054.

Batt, S. L., B. M. Charalambous, T. D. McHugh, S. Martin & S. H. Gillespie. 2005, "Novel PCR-restriction fragment length polymorphism method for determining serotypes or serogroups of *Streptococcus pneumoniae* isolates", *J. Clin. Microbiol.*, vol. 43, no. 6, 2656-2661.

Batt, S. L., B. M. Charalambous, A. W. Solomon, C. Knirsch, P. A. Massae, S. Safari, N. E. Sam, D. Everett, D. C. W. Mabey & S. H. Gillespie. 2003, "Impact of azithromycin administration for trachoma control on the carriage of antibiotic-resistant *Streptococcus pneumoniae*", *Antimicrob. Agents Chemother.*, vol. 47, no. 9, 2765-2769.

Bender, M., R. Cartee & J. Yother. 2003, "Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*", *J. Bacteriol.*, vol. 185, no. 20, 6057-6066.

Bender, M. & J. Yother. 2001, "CpsB is a modulator of capsule-associated tyrosine kinase activity in *Streptococcus pneumoniae*", *J. Biol. Chem.*, vol. 276, no. 51, 47966-47974.

Bentley, S. D., D. M. Aanensen, A. Mavroidi, D. Saunders, E. Rabbinowitsch, M. Collins, K. Donohoe, D. Harris, L. Murphy, M. A. Quail, G. Samuel, I. C. Skovsted, M. S. Kaltoft, B. Barrell, P. R. Reeves, J. Parkhill & B. G. Spratt. 2006, "Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes.", *PLoS Genet.*, vol. 2, no. 3, e31.

Bergeron, Y., N. Ouellet, A. Deslauriers, M. Simard, M. Olivier & M. Bergeron. 1998, "Cytokine kinetics and other host factors in response to pneumococcal pulmonary infection in mice", *Infect. Immun.*, vol. 66, no. 3, 912.

Berrón, S., A. Fenoll, M. Ortega, N. Arellano & J. Casal. 2005, "Analysis of the genetic structure of nontypeable pneumococcal strains isolated from conjunctiva", *J. Clin. Microbiol.*, vol. 43, no. 4, 1694-1698.

Berry, A. & J. Paton. 1996, "Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 64, no. 12, 5255-5262.

Berry, A. M., R. A. Lock, D. Hansman & J. C. Paton. 1989, "Contribution of autolysin to virulence of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 57, no. 8, 2324-2330.

Berry, A. M., R. A. Lock, S. M. Thomas, D. P. Rajan, D. Hansman & J. C. Paton. 1994, "Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*", *Infect. Immun.*, vol. 62, no. 3, 1101-1108.

Berry, A. M. & J. C. Paton. 2000, "Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins", *Infect. Immun.*, vol. 68, no. 1, 133-140.

Berry, A. M., J. C. Paton & D. Hansman. 1992, "Effect of insertional inactivation of the genes encoding pneumolysin and autolysin on the virulence of *Streptococcus pneumoniae* type 3.", *Microb. Pathog.*, vol. 12, no. 2, 87-93.

Billal, D. S., M. Hotomi, M. Suzumoto, K. Yamauchi, J. Arai, T. Katsurahara, S. Moriyama, K. Fujihara & N. Yamanaka. 2008, "Determination of pneumococcal serotypes/genotypes in nasopharyngeal secretions of otitis media children by multiplex PCR", *Eur. J. Pediatr.*, vol. 167, no. 4, 401-407.

Billal, D. S., M. Hotomi, S. Tasnim, K. Fujihara & N. Yamanaka. 2006, "Evaluation of serotypes of *Streptococcus pneumoniae* isolated from otitis media patients by multiplex polymerase chain reaction", *ORL*, vol. 68, no. 3, 135-138.

Black, S., H. Shinefield, R. Baxter, R. Austrian, L. Bracken, J. Hansen, E. Lewis & B. Fireman. 2004, "Postlicensure surveillance for pneumococcal invasive Disease after use of heptavalent pneumococcal conjugate vaccine in Northern California Kaiser Permanente", *Pediatr. Infect. Dis. J.*, vol. 23, no. 6, 485-489.

Black, S., H. Shinefield, B. Fireman, E. Lewis, P. Ray, J. R. Hansen, L. Elvin, K. M. Ensor, J. Hackell, G. Siber, F. Malinoski, D. Madore, I. Chang, R. Kohberger, W. Watson, R. Austrian & K. Edwards. 2000, "Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group", *Pediatr. Infect. Dis. J.*, vol. 19, no. 3, 187-195.

Black, S. B., H. R. Shinefield, S. Ling, J. Hansen, B. Fireman, D. Spring, J. Noyes, E. Lewis, P. Ray, J. Lee & J. Hackell. 2002, "Effectiveness of heptavalent pneumococcal conjugate vaccine in children younger than five years of age for prevention of pneumonia.", *Pediatr. Infect. Dis. J.*, vol. 21, no. 9, 810-815.

Bogaert, D., R. de Groot & P. Hermans. 2004, "Streptococcus pneumoniae colonisation: the key to pneumococcal disease", *Lancet Infect. Dis.,* vol. 4, no. 3, 144-154.

Bogaert, D., M. Engelen, A. Timmers-Reker, K. Elzenaar, P. Peerbooms, R. Coutinho, R. de Groot & P. Hermans. 2001, "Pneumococcal carriage in children in The Netherlands: a molecular epidemiological study", *J. Clin. Microbiol.,* vol. 39, no. 9, 3316-3320.

Bogaert, D., M. Sluijter, N. L.-d. Toom, T. J. Mitchell, W. H. F. Goessens, S. C. Clarke, R. de Groot & P. W. M. Hermans. 2006, "Dynamics of pneumococcal colonization in healthy Dutch children.", *Microbiology.,* vol. 152, no. 2, 377-385.

Bogaert, D., A. Van Belkum, M. Sluijter, A. Luijendijk, R. de Groot, H. Rümke, H. Verbrugh & P. Hermans. 2004a, "Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children", *Lancet,* vol. 363, no. 9424, 1871-1872.

Bogaert, D., R. Veenhoven, M. Sluijter, E. Sanders, R. Groot & P. Hermans. 2004b, "Colony blot assay: a useful method to detect multiple pneumococcal serotypes within clinical specimens", *FEMS Immun. Med. Microbiol.,* vol. 41, no. 3, 259-264.

Bootsma, H. J., M. Egmont-Petersen & P. W. M. Hermans. 2007, "Analysis of the in vitro transcriptional response of human pharyngeal epithelial cells to adherent *Streptococcus pneumoniae*: evidence for a distinct response to encapsulated strains", *Infect. Immun.,* vol. 75, no. 11, 5489-5499.

Borek, A. P., D. C. Dressel, J. Hussong & L. R. Peterson. 1997, "Evolving clinical problems with *Streptococcus pneumoniae*: increasing resistance to antimicrobial agents, and failure of traditional optochin identification in Chicago, Illinois, between 1993 and 1996.", *Diagn. Microbiol. Infect. Dis.,* vol. 29, no. 4, 209-214.

Bratcher, P. E., K.-H. Kim, J. H. Kang, J. Y. Hong & M. H. Nahm. 2010, "Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical and serological characterization", *Microbiology,* vol. 156, no. 2, 555-560.

Bratcher, P. E., I. H. Park, M. B. Oliver, M. Hortal, R. Camilli, S. K. Hollingshead, T. Camou & M. H. Nahm. 2011, "Evolution of the capsular gene locus of *Streptococcus pneumoniae* serogroup 6", *Microbiology,* vol. 157, no. 1, 189-198.

Braun, J. S., R. Novak, G. Gao, P. J. Murray & J. L. Shenep. 1999, "Pneumolysin, a protein toxin of *Streptococcus pneumoniae*, induces nitric oxide production from macrophages", *Infect. Immun.,* vol. 67, no. 8, 3750-3756.

Briles, D. E., E. Ades, J. C. Paton, J. S. Sampson, G. M. Carlone, R. C. Huebner, A. Virolainen, E. Swiatlo & S. K. Hollingshead. 2000a, "Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 68, no. 2, 796-800.

Briles, D. E., S. K. Hollingshead, J. King, A. Swift, P. A. Braun, M. K. Park, L. M. Ferguson, M. H. Nahm & G. S. Nabors. 2000b, "Immunization of humans with recombinant pneumococcal surface protein A (rPspA) elicits antibodies that passively protect mice from fatal infection with *Streptococcus pneumoniae* bearing heterologous PspA.", *J. Infect. Dis.,* vol. 182, no. 6, 1694-1701.

Briles, D. E., S. K. Hollingshead, J. C. Paton, E. W. Ades, L. Novak, F. W. van Ginkel & W. H. Benjamin. 2003, "Immunizations with pneumococcal surface protein A and pneumolysin are protective against pneumonia in a murine model of pulmonary infection with *Streptococcus pneumoniae*.", *J. Infect. Dis.,* vol. 188, no. 3, 339-348.

Brito, D., M. Ramirez & H. De Lencastre. 2003, "Serotyping *Streptococcus pneumoniae* by multiplex PCR", *J. Clin. Microbiol.,* vol. 41, no. 6, 2378-2384.

Brochet, M., E. Couve, M. Zouine, T. Vallaeys, C. Rusniok, M. C. Lamy, C. Buchrieser, P. Trieu-Cuot, F. Kunst, C. Poyart & P. Glaser. 2006, "Genomic diversity and evolution within the species *Streptococcus agalactiae*", *Microbes Infect,* vol. 8, no. 5, 1227-1243.

Bronsdon, M., K. O'Brien, R. Facklam, C. Whitney, B. Schwartz & G. Carlone. 2004, "Immunoblot method to detect *Streptococcus pneumoniae* and identify multiple serotypes from nasopharyngeal secretions", *J. Clin. Microbiol.,* vol. 42, no. 4, 1596-1600.

Brooks-Walter, A., D. E. Briles & S. K. Hollingshead. 1999, "The *pspC* gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia.", *Infect. Immun.,* vol. 67, no. 12, 6533-6542.

Brown, J., A. Ogunniyi, M. Woodrow, D. Holden & J. Paton. 2001, "Immunization with components of two iron uptake ABC transporters protects mice against systemic *Streptococcus pneumoniae* infection", *Infect. Immun.*, vol. 69, no. 11, 6702.

Brown, J. S., S. M. Gilliland & D. W. Holden. 2001, "A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence", *Mol. Microbiol.*, vol. 40, no. 3, 572-585.

Brown, J. S., S. M. Gilliland, J. Ruiz-Albert & D. W. Holden. 2002, "Characterization of pit, a *Streptococcus pneumoniae* iron uptake ABC transporter", *Infect. Immun.*, vol. 70, no. 8, 4389-4398.

Browne, K., J. Miegel & K. D. Stottmeier. 1984, "Detection of pneumococci in blood cultures by latex agglutination", *J. Clin. Microbiol.*, vol. 19, no. 5, 649-650.

Brückner, R., M. Nuhn, P. Reichmann, B. Weber & R. Hakenbeck. 2004, "Mosaic genes and mosaic chromosomes-genomic variation in *Streptococcus pneumoniae*", *Int. J. Med. Microbiol.*, vol. 294, no. 2-3, 157-168.

Brueggemann, A. & B. Spratt. 2003, "Geographic distribution and clonal diversity of Streptococcus pneumoniae serotype 1 isolates", *J. Clin. Microbiol.*, vol. 41, no. 11, 4966-4970.

Brueggemann, A. B., D. T. Griffiths, E. Meats, T. Peto, D. W. Crook & B. G. Spratt. 2003, "Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential", *J. Infect. Dis.*, vol. 187, no. 9, 1424-1432.

Brueggemann, A. B., R. Pai, D. W. Crook & B. Beall. 2007, "Vaccine escape recombinants emerge after pneumococcal vaccination in the United States", *PLoS Pathog.*, vol. 3, no. 11, e168.

Brugger, S. D., P. Frey, S. Aebi, J. Hinds & K. Mühlemann. 2010, "Multiple colonization with *S. pneumoniae* before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine", *PLoS ONE*, vol. 5, no. 7, e11638.

Brugger, S. D., L. J. Hathaway & K. Mühlemann. 2009, "Detection of *Streptococcus pneumoniae* strain cocolonization in the nasopharynx", *J. Clin. Microbiol.*, vol. 47, no. 6, 1750-1756.

Bryant, K., S. Block, S. Baker, W. Gruber & D. Scott. 2010, "Safety and immunogenicity of a 13-valent pneumococcal conjugate vaccine", *Pediatrics,* vol. 125, no. 5, 866-875.

Buck, J., C. Lexau, M. Shapiro, A. Glennen, D. Boxrud, B. Koziol, C. Whitney, B. Beall, R. Danila & R. Lynfield. 2006, "A community outbreak of conjunctivitis caused by nontypeable *Streptococcus pneumoniae* in Minnesota", *Pediatr. Infect. Dis. J.,* vol. 25, no. 10, 906-911.

Butler, J. C. (2004). Epidemiology of Pneumococcal Disease. The Pneumococcus. E. Tuomanen, A. M. Mitchell, D. A. Morrison and B. Spratt. Washington, D.C., ASM Press**:** 148-168.

Byrne, J. P., J. K. Morona, J. C. Paton & R. Morona. 2011, "Identification of *Streptococcus pneumoniae* Cps2C residues that affect capsular polysaccharide polymerization, cell wall ligation, and Cps2D phosphorylation", *J. Bacteriol.,* vol. 193, no. 9, 2341-2346.

Calix, J. J. & M. H. Nahm. 2010, "A new pneumococcal serotype, 11E, has a variably inactivated *wcjE* gene", *J. Infect. Dis.,* vol. 202, no. 1, 29-38.

Camilli, R., M. Del Grosso, F. Iannelli & A. Pantosti. 2008, "New genetic element carrying the erythromycin resistance determinant *erm(TR)* in *Streptococcus pneumoniae*", *Antimicrob. Agents Chemother.,* vol. 52, no. 2, 619-625.

Camilli, R., A. Pantosti & L. Baldassarri. 2011, "Contribution of serotype and genetic background to biofilm formation by *Streptococcus pneumoniae*", *Eur. J. Clin. Microbiol. Infect. Dis.,* vol. 30, no. 1, 97-102.

Campbell, E. A., S. Y. Choi & H. R. Masure. 1998, "A competence regulon in *Streptococcus pneumoniae* revealed by genomic analysis", *Mol. Microbiol.,* vol. 27, no. 5, 929-939.

Canvin, J. R., A. P. Marvin, M. Sivakumaran, J. C. Paton, G. J. Boulnois, P. W. Andrew & T. J. Mitchell. 1995, "The role of pneumolysin and autolysin in the pathology of pneumonia and septicemia in mice infected with a type 2 Pneumococcus", *J. Infect. Dis.,* vol. 172, no. 1, 119-123.

Capeding, M. R., H. Nohynek, L. T. Sombrero, L. G. Pascual, E. S. Sunico, G. A. Esparar, E. Esko, M. Leinonen & P. Ruutu. 1995, "Evaluation of sampling sites for detection of upper respiratory tract carriage of *Streptococcus pneumoniae* and *Haemophilus influenzae* among healthy Filipino infants.", *J. Clin. Microbiol.,* vol. 33, no. 11, 3077-3079.

Cardozo, D. M., C. M. Nascimento-Carvalho, A.-L. S. S. Andrade, A. M. Silvany-Neto, C. H. C. Daltro, M.-A. S. Brandão, A. P. Brandao & M. C. C. Brandileone. 2008, "Prevalence and risk factors for nasopharyngeal carriage of *Streptococcus pneumoniae* among adolescents.", *J. Med. Microbiol.*, vol. 57, no. 2, 185-189.

Carrolo, M., F. R. Pinto, J. Melo-Cristino & M. Ramirez. 2009, "Pherotypes are driving genetic differentiation within *Streptococcus pneumoniae*", *BMC Microbiology*, vol. 9, no. 191.

Cartee, R. T., W. T. Forsee, M. H. Bender, K. D. Ambrose & J. Yother. 2005, "CpsE from type 2 *Streptococcus pneumoniae* catalyzes the reversible addition of glucose-1-phosphate to a polyprenyl phosphate acceptor, initiating type 2 capsule repeat unit formation", *J. Bacteriol.*, vol. 187, no. 21, 7425-7433.

Cartee, R. T., W. T. Forsee, J. S. Schutzbach & J. Yother. 2000, "Mechanism of type 3 capsular polysaccharide synthesis in *Streptococcus pneumoniae*", *J. Biol. Chem.*, vol. 275, no. 6, 3907-3914.

Carter, R. J. F. 2006, "Infants too young to receive pneumococcal conjugate vaccine benefit from herd immunity", *Thorax*, vol. 61, no. 7, 610-610.

Carvalho, M., F. C. Pimenta, R. E. Gertz, H. H. Joshi, A. A. Trujillo, L. E. Keys, J. Findley, I. S. Moura, I. H. Park, S. K. Hollingshead, T. Pilishvili, C. G. Whitney, M. H. Nahm & B. W. Beall. 2009, "PCR-based quantitation and clonal diversity of the current prevalent invasive serogroup 6 pneumococcal serotype, 6C, in the United States in 1999 and 2006 to 2007", *J. Clin. Microbiol.*, vol. 47, no. 3, 554-559.

Carvalho, M., A. Steigerwalt, T. Thompson, D. Jackson & R. Facklam. 2003, "Confirmation of nontypeable *Streptococcus pneumoniae*-like organisms isolated from outbreaks of epidemic conjunctivitis as *Streptococcus pneumoniae*", *J. Clin. Microbiol.*, vol. 41, no. 9, 4415-4417

Carvalho, M. d. G., F. C. Pimenta, D. Jackson, A. Roundtree, Y. Ahmad, E. V. Millar, K. L. O'Brien, C. G. Whitney, A. L. Cohen & B. W. Beall. 2010, "Revisiting pneumococcal carriage by use of broth enrichment and PCR techniques for enhanced detection of carriage and serotypes", *J. Clin. Microbiol.*, vol. 48, no. 5, 1611-1618.

Carvalho, M. d. G. S., M. L. Tondella, K. McCaustland, L. Weidlich, L. McGee, L. W. Mayer, A. Steigerwalt, M. Whaley, R. R. Facklam, B. Fields, G. Carlone, E. W. Ades, R. Dagan & J. S. Sampson. 2007, "Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA", *J. Clin. Microbiol.*, vol. 45, no. 8, 2460-2466.

Charalambous, B., S. Batt, A. Peek, H. Mwerinde, N. Sam & S. Gillespie. 2003, "Quantitative validation of media for transportation and storage of *Streptococcus pneumoniae*", *J. Clin. Microbiol.*, vol. 41, no. 12, 5551-5556.

Charalambous, B. M., N. M. Oriyo & S. H. Gillespie. 2008, "How valid is single-colony isolation for surveillance of *Streptococcus pneumoniae* carriage?", *J. Clin. Microbiol.*, vol. 46, no. 7, 2467-2468.

Chen, C.-J., Y.-C. Huang, L.-H. Su & T.-Y. Lin. 2007, "Nasal carriage of *Streptococcus pneumoniae* in healthy children and adults in northern Taiwan", *Diagn. Microbiol. Infect. Dis.*, vol. 59, no. 3, 265-269.

Chi, F., M. Leider, F. Leendertz, C. Bergmann, C. Boesch, S. Schenk, G. Pauli, H. Ellerbrok & R. Hakenbeck. 2007, "New *Streptococcus pneumoniae* clones in deceased wild chimpanzees", *J. Bacteriol.*, vol. 189, no. 16, 6085-6088.

Cieslewicz, M. J., D. L. Kasper, Y. Wang & M. R. Wessels. 2001, "Functional analysis in type Ia group B *Streptococcus* of a cluster of genes involved in extracellular polysaccharide production by diverse species of streptococci.", *J. Biol. Chem.*, vol. 276, no. 1, 139-146.

Claverys, J.-P. & L. S. Havarstein. 2002, "Extracellular-peptide control of competence for genetic transformation in *Streptococcus pneumoniae*", *Front. Biosci.*, vol. 7, no. d1798-1814.

Clinical and Laboratory Standards Institute. Performance Standards for Antimicrobial Susceptibility Testing: Eighteenth Informational Supplement. 2008.

Cochetti, I., M. Vecchi, M. Mingoia, E. Tili, M. R. Catania, A. Manzin, P. E. Varaldo & M. P. Montanari. 2005, "Molecular characterization of pneumococci with efflux-mediated erythromycin resistance and identification of a novel *mef* gene subclass, *mef(I)*.", *Antimicrob. Agents Chemother.*, vol. 49, no. 12, 4999-5006.

Coffey, T. J., M. Daniels, M. C. Enright & B. G. Spratt. 1999, "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of *Streptococcus pneumoniae* arose by large recombinational replacements of the *cpsA-pbp1a* region", *Microbiology*, vol. 145, no. 8, 2023-2031.

Coffey, T. J., C. G. Dowson, M. Daniels, J. Zhou, C. Martin, B. G. Spratt & J. M. Musser. 1991, "Horizontal transfer of multiple penicillin-binding protein genes, and capsular biosynthetic genes, in natural populations of *Streptococcus pneumoniae*", *Mol. Microbiol.,* vol. 5, no. 9, 2255-2260.

Coffey, T. J., M. C. Enright, M. Daniels, J. K. Morona, R. Morona, W. Hryniewicz, J. C. Paton & B. G. Spratt. 1998, "Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*", *Mol. Microbiol.,* vol. 27, no. 1, 73-83.

Cohen, R., C. Levey, E. Bonnett, S. Grondin, V. Desvignes, A. Lecuyer, B. Fritzell & E. Varon. 2010, "Dynamic of pneumococcal nasopharyngeal carriagein children with acute otitis media following PCV7 introduction in France", *Vaccine*, vol. 28, no. 37, 6114-6121.

Cooke, B., A. Smith, M. Diggle, K. Lamb, C. Robertson, D. Inverarity, J. Jefferies, G. Edwards, T. Mitchell, S. Clarke & J. McMenamin. 2010, "Antibiotic resistance in invasive *Streptococcus pneumoniae* isolates identified in Scotland between 1999 and 2007.", *J. Med. Microbiol.,* vol. 59, no. 10, 1212-1218.

Cooper, D., X. Yu, M. Sidhu, M. H. Nahm, P. Fernsten & K. U. Jansen. 2011, "The 13-valent pneumococcal conjugate vaccine (PCV13) elicits cross-functional opsonophagocytic killing responses in humans to *Streptococcus pneumoniae* serotypes 6C and 7A", *Vaccine,* vol. 29, no. 41, 7207-7211.

Cornejo, O. E., L. McGee & D. E. Rozen. 2010, "Polymorphic competence peptides do not restrict recombination in *Streptococcus pneumoniae*", *Mol. Biol. Evol.,* vol. 27, no. 3, 694-702.

Croucher, N. J., S. R. Harris, C. Fraser, M. A. Quail, J. Burton, M. van der Linden, L. McGee, A. von Gottberg, J. H. Song, K. S. Ko, B. Pichon, S. Baker, C. M. Parry, L. M. Lambertsen, D. Shahinas, D. R. Pillai, T. J. Mitchell, G. Dougan, A. Tomasz, K. P. Klugman, J. Parkhill, W. P. Hanage & S. D. Bentley. 2011, "Rapid pneumococcal evolution in response to clinical interventions", *Science,* vol. 331, no. 6016, 430-434.

Crum, N. F., C. P. Barrozo, F. A. Chapman, M. A. K. Ryan & K. L. Russell. 2004, "An outbreak of conjunctivitis due to a novel unencapsulated *Streptococcus pneumoniae* among military trainees", *Clin. Infect. Dis.,* vol. 39, no. 8, 1148-1154.

Crump, J. A., H. O. Ramadhani, A. B. Morrissey, L. J. Msuya, L.-Y. Yang, S.-C. Chow, S. C. Morpeth, H. Reyburn, B. N. Njau, A. V. Shaw, H. C. Diefenthal, J. A. Bartlett, J. F. Shao, W. Schimana, C. K. Cunningham & G. D. Kinabo. 2011a, "Invasive bacterial and fungal infections among hospitalized HIV-infected and HIV-uninfected children and infants in northern Tanzania", *Trop. Med. Int. Health.*, vol. 16, no. 7, 830-837.

Crump, J. A., H. O. Ramadhani, A. B. Morrissey, W. Saganda, M. S. Mwako, L.-Y. Yang, S.-C. Chow, S. C. Morpeth, H. Reyburn, B. N. Njau, A. V. Shaw, H. C. Diefenthal, J. F. Shao, J. A. Bartlett & V. P. Maro. 2011b, "Invasive bacterial and fungal infections among hospitalized HIV-infected and HIV-uninfected adults and adolescents in northern Tanzania", *Clin. Infect. Dis.*, vol. 52, no. 3, 341-348.

Cundell, D. R., N. P. Gerard, C. Gerard, I. Idanpaan-Heikkila & E. I. Tuomanen. 1995, "*Streptococcus pneumoniae* anchor to activated human cells by the receptor for platelet-activating factor.", *Nature*, vol. 377, no. 6548, 435-438.

Curns, A. T., R. C. Holman, D. K. Shay, J. E. Cheek, S. F. Kaufman, R. J. Singleton & L. J. Anderson. 2002, "Outpatient and hospital visits associated with otitis media among American Indian and Alaska native children younger than 5 years.", *Pediatrics*, vol. 109, no. 3, E41-41.

Cutts, F. T., S. M. A. Zaman, G. Enwere, S. Jaffar, O. S. Levine, J. B. Okoko, C. Oluwalana, A. Vaughan, S. K. Obaro, A. Leach, K. P. McAdam, E. Biney, M. Saaka, U. Onwuchekwa, F. Yallop, N. F. Pierce, B. M. Greenwood, R. A. Adegbola & G. P. V. T. Group. 2005, "Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial.", *Lancet*, vol. 365, no. 9465, 1139-1146.

Dagan, R., N. Givon-Lavi, O. Zamir & D. Fraser. 2003, "Effect of a nonavalent conjugate vaccine on carriage of antibiotic-resistant *Streptococcus pneumoniae* in day-care centers", *Pediatr. Infect. Dis. J.*, vol. 22, no. 6, 532-540.

Dagan, R., N. Givon-Lavi, O. Zamir, M. Sikuler-Cohen, L. Guy, J. Janco, P. Yagupsky & D. Fraser. 2002, "Reduction of nasopharyngeal carriage of *Streptococcus pneumoniae* after administration of a 9-valent pneumococcal conjugate vaccine to toddlers attending day care centers", *J. Infect. Dis.*, vol. 185, no. 7, 927-936.

Dagan, R., R. Melamed, M. Muallem, L. Piglansky & P. Yagupsky. 1996, "Nasopharyngeal colonization in southern Israel with antibiotic-resistant pneumococci during the first 2 years of life: relation to serotypes likely to be

included in pneumococcal conjugate vaccines.", *J. Infect. Dis.*, vol. 174, no. 6, 1352-1355.

Dagerhamn, J., C. Blomberg, S. Browall, K. Sjostrom, E. Morfeldt & B. Henriques-Normark. 2008, "Determination of accessory gene patterns predicts the same relatedness among strains of *Streptococcus pneumoniae* as sequencing of housekeeping genes does and represents a novel approach in molecular epidemiology", *J. Clin. Microbiol.*, vol. 46, no. 3, 863-868.

Dagkessamanskaia, A., M. Moscoso, V. Hénard, S. Guiral, K. Overweg, M. Reuter, B. Martin, J. Wells & J.-P. Claverys. 2004, "Interconnection of competence, stress and CiaR regulons in *Streptococcus pneumoniae*: competence triggers stationary phase autolysis of *ciaR* mutant cells", *Mol. Microbiol.*, vol. 51, no. 4, 1071-1086.

Darboe, M., A. Fulford, O. Secka & A. Prentice. 2010, "The dynamics of nasopharyngeal *Streptococcus pneumoniae* carriage among rural Gambian mother-infant pairs", *BMC Infect. Dis.*, vol. 10, no. 1, 195.

Dave, S., A. Brooks-Walter, M. K. Pangburn & L. S. McDaniel. 2001, "PspC, a pneumococcal surface protein, binds human factor H", *Infect. Immun.*, vol. 69, no. 5, 3435-3437.

Davie, J. J., J. Earl, S. P. W. de Vries, A. Ahmed, F. Z. Hu, H. J. Bootsma, K. Stol, P. W. M. Hermans, R. M. Wadowsky, G. D. Ehrlich, J. P. Hays & A. A. Campagnari. 2011, "Comparative analysis and supragenome modeling of twelve *Moraxella catarrhalis* clinical isolates", *BMC Genomics*, vol. 12, no. 70

Dawid, S., A. Roche & J. Weiser. 2007, "The *blp* bacteriocins of *Streptococcus pneumoniae* mediate intraspecies competition both *in vitro* and *in vivo*", *Infect. Immun.*, vol. 75, no. 1, 443-451

De Las Rivas, B., J. Garcia, R. Lopez & P. Garcia. 2002, "Purification and polar localization of pneumococcal LytB, a putative endo-{beta}-*N*-acetylglucosaminidase: the chain-dispersing murein hydrolase", *J. Bacteriol.*, vol. 184, no. 18, 4988-5000

Denapaite, D., R. Brückner, R. Hakenbeck & W. Vollmer. 2012, "Biosynthesis of teichoic acids in *Streptococcus pneumoniae* and closely related species: lessons from genomes.", *Microb. Drug Resist.*, doi:10.1089/mdr.2012.0026.

Denapaite, D., R. Bruckner, M. Nuhn, P. Reichmann, B. Henrich, P. Maurer, Y. Schahle, P. Selbmann, W. Zimmermann, R. Wambutt & R. Hakenbeck. 2010, "The genome of *Streptococcus mitis* B6--what is a commensal?", *PLoS ONE,* vol. 5, no. 2, e9426.

Denapaite, D. & R. Hakenbeck. 2011, "A new variant of the capsule 3 cluster occurs in *Streptococcus pneumoniae* from deceased wild chimpanzees", *PLoS ONE,* vol. 6, no. 9, e25119.

Di Guilmi, A. M. & A. Dessen. 2002, "New approaches towards the identification of antibiotic and vaccine targets in *Streptococcus pneumoniae*.", *EMBO Rep.,* vol. 3, no. 8, 728-734.

Dias, C., L. Teixeira, M. Carvalho & B. Beall. 2007, "Sequential multiplex PCR for determining capsular serotypes of pneumococci recovered from Brazilian children", *J. Med. Microbiol.,* vol. 56, no. 9, 1185-1188.

Diebel, R. H. & W. J. Seeley (1974). Streptococcaceae. Bergey's Manual of Determinative Bacteriology, 8th edn. R. E. Buchanan and N. E. Gibbons. Baltimore, The Williams & Wilkins Co.**:** 490-509.

Dillard, J. & J. Yother. 1994, "Genetic and molecular characterization of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* type 3", *Mol. Microbiol.,* vol. 12, no. 6, 959-972.

Dillard, J. P., M. W. Vandersea & J. Yother. 1995, "Characterization of the cassette containing genes for type 3 capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*", *J. Exp. Med.,* vol. 181, no. 3, 973-983.

Dilworth, J. A., P. Stewart, J. M. Gwaltney, J. O. Hendley & M. A. Sande. 1975, "Methods to improve detection of pneumococci in respiratory secretions", *J. Clin. Microbiol.,* vol. 2, no. 5, 453-455.

Ding, F., P. Tang, M. Hsu, P. Cui, S. Hu, J. Yu & C. Chiu. 2009, "Genome evolution driven by host adaptations results in a more virulent and antimicrobial-resistant *Streptococcus pneumoniae* serotype 14", *BMC Genomics,* vol. 10, no. 1, 158.

Dirix, G., P. Monsieurs, K. Marchal, J. Vanderleyden & J. Michiels. 2004, "Screening genomes of Gram-positive bacteria for double-glycine-motif-containing peptides", *Microbiol.,* vol. 150, no. 5, 1121-1126.

Domenech, M., E. García & M. Moscoso. 2009, "Versatility of the capsular genes during biofilm formation by *Streptococcus pneumoniae*", *Environ. Microbiol.,* vol. 11, no. 10, 2542-2555.

Domenech, M., E. García & M. Moscoso. 2011, "*In vitro* destruction of *Streptococcus pneumoniae* biofilms with bacterial and phage peptidoglycan hydrolases", *Antimicrob. Agents Chemother.,* vol. 55, no. 9, 4144-4148.

Donati, C., N. Hiller, H. Tettelin, A. Muzzi, N. Croucher, S. Angiuoli, M. Oggioni, J. Hotopp, F. Hu & D. Riley. 2010, "Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species", *Genome Biol.,* vol. 11, no. 10, R107.

Donkor, E. S., C. J. Bishop, M. Antonio, B. Wren & W. P. Hanage. 2011, "High levels of recombination among Streptococcus pneumoniae isolates from the Gambia", *mBio,* vol. 2, no. 3, e00040-00011.

Donlan, R. M. 2000, "Role of biofilms in antimicrobial resistance", *ASAIO J.,* vol. 46, no. 6, S47-52.

Dopazo, J., A. Mendoza, J. Herrero, F. Caldara, Y. Humbert, L. Friedli, M. Guerrier, E. Grand-Schenk, C. Gandin, M. de Francesco, A. Polissi, G. Buell, G. Feger, E. Garcia, M. Peitsch, and J. F. Garcia-Bustos. 2001, "Anotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate", *Microb. Drug Resist.,* vol. 7, no. 2, 99-125.

Douglas, R. M. & H. B. Miles. 1984, "Vaccination against *Streptococcus pneumoniae* in childhood: lack of demonstrable benefit in young Australian children", *J. Infect. Dis.,* vol. 149, no. 6, 861-869.

Dowling, J. N., P. R. Sheehe & H. A. Feldman. 1971, "Pharyngeal pneumococcal acquisitions in 'normal' families: a longitudinal study", *J. Infect. Dis.,* vol. 124, no. 1, 9-17.

Dowson, C., T. Coffey, C. Kell & R. Whiley. 1993, "Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of *Streptococcus mitis* in the formation of a low affinity PBP2B in *S. pneumoniae*", *Mol. Microbiol.,* vol. 9, no. 3, 635-643.

Dowson, C. G., V. Barcus, S. King, P. Pickerill, A. Whatmore & M. Yeo. 1997, "Horizontal gene transfer and the evolution of resistance and virulence determinants in *Streptococcus*", *Soc. Appl. Bacteriol. Symp. Ser.,* vol. 26, no. 42S-51S.

Dowson, C. G., A. Hutchison, J. A. Brannigan, R. C. George, D. Hansman, J. Liñares, A. Tomasz, J. M. Smith & B. G. Spratt. 1989, "Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 86, no. 22, 8842-8846.

Dowson, C. G., A. Hutchison, N. Woodford, A. P. Johnson, R. C. George & B. G. Spratt. 1990, "Penicillin-resistant viridans streptococci have obtained altered penicillin-binding protein genes from penicillin-resistant strains of *Streptococcus pneumoniae*", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 87, no. 15, 5858-5862.

Dudley, S., K. Ashe, B. Winther & J. O. Hendley. 2001, "Bacterial pathogens of otitis media and sinusitis: detection in the nasopharynx with selective agar media", *J. Lab. Clin. Med.*, vol. 138, no. 5, 338-342.

Dunne, E. M., J. Manning, F. M. Russell, R. M. Robins-Browne, E. K. Mulholland & C. Satzke. 2012, "Effect of pneumococcal vaccination on nasopharyngeal carriage of *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Moraxella catarrhalis*, and *Staphylococcus aureus* in Fijian Children", *J. Clin. Microbiol.*, vol. 50, no. 3, 1034-1038.

Ehrlich, G. D., F. Z. Hu, K. Shen, P. Stoodley & J. C. Post. 2005, "Bacterial plurality as a general mechanism driving persistence in chronic infections", *Clin. Orthop. Relat. Res.*, no. 437, 20-24.

Ehrlich, G. D., R. Veeh, X. Wang, J. W. Costerton, J. D. Hayes, F. Z. Hu, B. J. Daigle, M. D. Ehrlich & J. C. Post. 2002, "Mucosal biofilm formation on middle-ear mucosa in the chinchilla model of otitis media", *JAMA*, vol. 287, no. 13, 1710-1715.

El Ahmer, O. R., S. D. Essery, A. T. Saadi, M. W. Raza, M. M. Ogilvie, D. M. Weir & C. C. Blackwell. 1999, "The effect of cigarette smoke on adherence of respiratory pathogens to buccal epithelial cells", *FEMS Immun. Med. Microbiol.*, vol. 23, no. 1, 27-36.

El Aila, N. A., S. Emler, T. Kaijalainen, T. De Baere, B. Saerens, E. Alkan, P. Deschaght, R. Verhelst & M. Vaneechoutte. 2010, "The development of a 16S rRNA gene based PCR for the identification of *Streptococcus pneumoniae* and comparison with four other species specific PCR assays", *BMC Infect. Dis.*, vol. 10, no. 104.

El-Refaie, M. & C. Dulake. 1975, "Counter-current immunoelectrophoresis for the diagnosis of pneumococcal chest infection", *J. Clin. Pathol.*, vol. 28, no. 10, 801-806.

Elberse, K., S. Witteveen, H. van der Heide, I. van de Pol, C. Schot, A. van der Ende, G. Berbers & L. Schouls. 2011, "Sequence diversity within the capsular genes of *Streptococcus pneumoniae* serogroup 6 and 19", *PLoS ONE,* vol. 6, no. 9, e25018.

Eldholm, V., O. Johnsborg, K. Haugen, H. S. Ohnstad & L. S. Håvarstein. 2009, "Fratricide in *Streptococcus pneumoniae*: contributions and role of the cell wall hydrolases CbpD, LytA and LytC", *Microbiology,* vol. 155, no. 7, 2223-2234.

Eldholm, V., O. Johnsborg, D. Straume, H. S. Ohnstad, K. H. Berg, J. A. Hermoso & L. S. Håvarstein. 2010, "Pneumococcal CbpD is a murein hydrolase that requires a dual cell envelope binding specificity to kill target cells during fratricide", *Mol. Microbiol.*, vol. 76, no. 4, 905-917.

Enright, M. & B. Spratt. 1999, "Extensive variation in the *ddl* gene of penicillin-resistant *Streptococcus pneumoniae* results from a hitchhiking effect driven by the penicillin-binding protein 2b gene.", *Mol. Biol. Evol.*, vol. 16, no. 12, 1687-1695.

Enright, M. C. & B. G. Spratt. 1998, "A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease", *Microbiology,* vol. 144, no. 11, 3049-3060.

Eskola, J., T. Kilpi, A. Palmu, J. Jokinen, J. Haapakoski, E. Herva, A. Takala, H. Käyhty, P. Karma, R. Kohberger, G. Siber, P. H. Mäkelä & F. O. M. S. Group. 2001, "Efficacy of a pneumococcal conjugate vaccine against acute otitis media", *N. Engl. J. Med.*, vol. 344, no. 6, 403-409.

Faden, H., L. Duffy, R. Wasielewski, J. Wolf, D. Krystofik & Y. Tung. 1997, "Relationship between nasopharyngeal colonization and the development of otitis media in children. Tonawanda/Williamsville Pediatrics", *J. Infect. Dis.*, vol. 175, no. 6, 1440-1445.

Fainstein, V., D. M. Musher & T. R. Cate. 1980, "Bacterial adherence to pharyngeal cells during viral infection", *J. Infect. Dis.*, vol. 141, no. 2, 172-176.

Feil, E. J., M. C. Enright & B. G. Spratt. 2000, "Estimating the relative contributions of mutation and recombination to clonal diversification: a comparison between *Neisseria meningitidis* and *Streptococcus pneumoniae*", *Res. Microbiol.*, vol. 151, no. 6, 465-469.

Feil, E. J., B. C. Li, D. M. Aanensen, W. P. Hanage & B. G. Spratt. 2004, "eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data", *J. Bacteriol.*, vol. 186, no. 5, 1518-1530.

Fenoll, A., L. Aguilar, M.-J. Gimenez, M.-D. Vicioso, O. Robledo, J.-J. Granizo & P. Coronel. 2010, "Susceptibility of recently collected Spanish pneumococci nonsusceptible to oral penicillin from serotypes not included in the 7-valent conjugate vaccine.", *Antimicrob. Agents and Chemother.*, vol. 54, no. 6, 2696-2698.

Fenoll, A., J. J. Granizo, L. Aguilar, M. J. Giménez, L. Aragoneses-Fenoll, G. Hanquet, J. Casal & D. Tarragó. 2009, "Temporal trends of invasive *Streptococcus pneumoniae* serotypes and antimicrobial resistance patterns in Spain from 1979 to 2007", *J. Clin. Microbiol.*, vol. 47, no. 4, 1012-1020.

Fenoll, A., I. Jado, D. Vicioso & J. Casal. 1997, "Dot blot assay for the serotyping of pneumococci.", *J. clin. Microbiol.*, vol. 35, no. 3, 764-766.

Fenoll, A., C. Martín Bourgon, R. Muñóz, D. Vicioso & J. Casal. 1991, "Serotype distribution and antimicrobial resistance of *Streptococcus pneumoniae* isolates causing systemic infections in Spain, 1979-1989", *Rev. Infect. Dis.*, vol. 13, no. 1, 56-60.

Filippidis, A. & K. N. Fountas. 2009, "Nasal lymphatics as a novel invasion and dissemination route of bacterial meningitis", *Med. Hypotheses*, vol. 72, no. 6, 694-697.

Finland, M. & W. D. Sutliff. 1931, "Specific cutaneous reactions and circulating antibodies in the course of lobar pneumonia: II. Cases treated with antipneumococcic sera", *J. Exp. Med.*, vol. 54, no. 5, 653-667.

Fireman, B., S. B. Black, H. R. Shinefield, J. Lee, E. Lewis & P. Ray. 2003, "Impact of the pneumococcal conjugate vaccine on otitis media", *Pediatr. Infect. Dis. J.*, vol. 22, no. 1, 10-16.

Fischer, W. 2000, "Phosphocholine of pneumococcal teichoic acids: role in bacterial physiology and pneumococcal infection", *Res. Microbiol.*, vol. 151, no. 6, 421-427.

Fitzgerald, J. R., D. E. Sturdevant, S. M. Mackie, S. R. Gill & J. M. Musser. 2001, "Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic.", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 98, no. 15, 8821-8826.

Flamaing, J., W. Peetermans, J. Vandeven & J. Verhaegen. 2010, "Pneumococcal colonization in older persons in a nonoutbreak setting", *J. Am. Geriatr. Soc.*, vol. 58, no. 2, 396-398.

Fossieck, B., Jr., R. Craig & P. Y. Paterson. 1973, "Counterimmunoelectrophoresis for rapid diagnosis of meningitis due to *Diplococcus pneumoniae*", *J. Infect. Dis*, vol. 127, no. 1, 106-109.

Freney, J., S. Bland, J. Etienne, M. Desmonceaux, J. M. Boeufgras & J. Fleurette. 1992, "Description and evaluation of the semiautomated 4-hour rapid ID 32 Strep method for identification of streptococci and members of related genera.", *J. Clin. Microbiol.*, vol. 30, no. 10, 2657-2661.

Freyer, D., R. Manz, A. Ziegenhorn, M. Weih, K. Angstwurm, W. D. Döcke, A. Meisel, R. R. Schumann, G. Schönfelder, U. Dirnagl & J. R. Weber. 1999, "Cerebral endothelial cells release TNF-alpha after stimulation with cell walls of *Streptococcus pneumoniae* and regulate inducible nitric oxide synthase and ICAM-1 expression via autocrine loops", *J. Immun.*, vol. 163, no. 8, 4308-4314.

García, P., M. P. González, E. García, R. López & J. L. García. 1999a, "LytB, a novel pneumococcal murein hydrolase essential for cell separation", *Mol. Microbiol.*, vol. 31, no. 4, 1275-1281.

García, P., M. Paz González, E. García, J. L. García & R. López. 1999b, "The molecular characterization of the first autolytic lysozyme of *Streptococcus pneumoniae* reveals evolutionary mobile domains", *Mol. Microbiol.*, vol. 33, no. 1, 128-138.

Garcia-Bustos, J. & A. Tomasz. 1990, "A biological price of antibiotic resistance: major changes in the peptidoglycan structure of penicillin-resistant pneumococci", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 87, no. 14, 5415-5419.

García-Castillo, M., M. I. Morosini, A. Valverde, F. Almaraz, F. Baquero, R. Cantón & R. del Campo. 2007, "Differences in biofilm development and antibiotic susceptibility among *Streptococcus pneumoniae* isolates from cystic fibrosis samples and blood cultures", *J. Antimicrob. Chemother.*, vol. 59, no. 2, 301-304.

Gay, K. & D. S. Stephens. 2001, "Structure and dissemination of a chromosomal insertion element encoding macrolide efflux in *Streptococcus pneumoniae*", *J. Infect. Dis.*, vol. 184, no. 1, 56-65.

Georgopapadakou, N. H. 1993, "Penicillin-binding proteins and bacterial resistance to beta-lactams.", *Antimicrob. Agents Chemother.*, vol. 37, no. 10, 2045-2053.

Gessner, B. D., J. E. Mueller & S. Yaro. 2010, "African meningitis belt pneumococcal disease epidemiology indicates a need for an effective serotype 1 containing vaccine, including for older children and adults", *BMC Infect. Dis.*, vol. 10, no. 22.

Gherna, R. L. (1981). Preservation. Manual of Methods for General Bacteriology. Washington, D.C., American Society for Microbiology**:** 208-217.

Giammarinaro, P. & J. C. Paton. 2002, "Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 70, no. 10, 5454-5461.

Gibson, L. & J. Khoury. 1986, "Storage and survival of bacteria by ultra-freeze", *Lett. Appl. Microbiol.*, vol. 3, no. 6, 127-129.

Gill, C., V. Mwanakasale, M. Fox, R. Chilengi, M. Tembo, M. Nsofwa, V. Chalwe, L. Mwananyanda, D. Mukwamataba & B. Malilwe. 2008a, "Effect of presumptive co-trimoxazole prophylaxis on pneumococcal colonization rates, seroepidemiology and antibiotic resistance in Zambian infants: a longitudinal cohort study", *Bull. World Health Organ.*, vol. 86, no. 12, 929-938.

Gill, C., V. Mwanakasale, M. Fox, R. Chilengi, M. Tembo, M. Nsofwa, V. Chalwe, L. Mwananyanda, D. Mukwamataba & B. Malilwe. 2008b, "Impact of human immunodeficiency virus infection on *Streptococcus pneumoniae* colonization and seroepidemiology among Zambian women", *J. Infect. Dis.*, vol. 197, no. 7, 1000-1005.

Gillespie, S. H. & I. Balakrishnan. 2000, "Pathogenesis of pneumococcal infection", *J. Med. Microbiol.*, vol. 49, no. 12, 1057-1067.

Givon-Lavi, N., R. Dagan, D. Fraser, P. Yagupsky & N. Porat. 1999, "Marked differences in pneumococcal carriage and resistance patterns between day care centers located within a small area.", *Clin. Infect. Dis.,* vol. 29, no. 5, 1274-1280.

Gladstone, R. A., J. M. Jefferies, S. N. Faust & S. C. Clarke. 2011a, "Continued control of pneumococcal disease in the UK - the impact of vaccination", *J. Med. Microbiol.,* vol. 60, no. 1, 1-8.

Gladstone, R. A., J. M. Jefferies, S. N. Faust & S. C. Clarke (2011b). Detection of pneumococcal carriage: a systematic view. Society for General Microbiology Spring Conference 2011. Harrogate, UK.

Global Alliance for Vaccines and Immunizations, (2012). "Pneumococcal Vaccine Support." 2011, from http://www.gavialliance.org/support/nvs/pneumococcal/.

Goldblatt, D., M. Hussain, N. Andrews, L. Ashton, C. Virta, A. Melegaro, R. Pebody, R. George, A. Soininen, J. Edmunds, N. Gay, H. Kayhty & E. Miller. 2005, "Antibody responses to nasopharyngeal carriage of *Streptococcus pneumoniae* in adults: a longitudinal household study", *J. Infect. Dis.,* vol. 192, no. 3, 387-393.

Golubchik, T., A. B. Brueggemann, T. Street, R. E. Gertz, C. C. A. Spencer, T. Ho, E. Giannoulatou, R. Link-Gelles, R. M. Harding, B. Beall, T. E. A. Peto, M. R. Moore, P. Donnelly, D. W. Crook & R. Bowden. 2012, "Pneumococcal genome sequencing tracks a vaccine escape variant formed through a multi-fragment recombination event.", *Nature Genet.,* vol. 44, no. 3, 352-355.

Gordon, S. B., G. R. Irving, R. A. Lawson, M. E. Lee & R. C. Read. 2000, "Intracellular trafficking and killing of *Streptococcus pneumoniae* by human alveolar macrophages are influenced by opsonins", *Infect. Immun.,* vol. 68, no. 4, 2286-2293.

Gosink, K. K., E. R. Mann, C. Guglielmo, E. I. Tuomanen & H. R. Masure. 2000, "Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 68, no. 10, 5690-5695.

Granizo, J. J., L. Aguilar, J. Casal, C. García-Rey, R. Dal-Ré & F. Baquero. 2000, "*Streptococcus pneumoniae* resistance to erythromycin and penicillin in relation to macrolide and beta-lactam consumption in Spain (1979-1997).", *J. Antimicrob. Chemother.,* vol. 46, no. 5, 767-773.

Gratten, M., K. Manning, J. Dixon, F. Morey, P. Torzillo, J. Hanna, J. Erlich, V. Asche & I. Riley. 1994, "Upper airway carriage by *Haemophilus influenzae* and *Streptococcus pneumoniae* in Australian aboriginal children hospitalised with acute lower respiratory infection", *Southeast Asian J Trop Med Public Health,* vol. 25, no. 1, 123-131.

Gratten, M., J. Montgomery, G. Gerega, H. Gratten, H. Siwi, A. Poli & G. Koki. 1989, "Multiple colonization of the upper respiratory tract of Papua New Guinea children with *Haemophilus influenzae* and *Streptococcus pneumoniae*", *Southeast Asian J. Trop. Med. Public Health,* vol. 20, no. 4, 501-509.

Grivea, I. N., A. G. Tsantouli, A. N. Michoula & G. A. Syrogiannopoulos. 2011, "Dynamics of *Streptococcus pneumoniae* nasopharyngeal carriage with high heptavalent pneumococcal conjugate vaccine coverage in Central Greece", *Vaccine,* vol. 8, no. 29, 8882-8887.

Gray, B. M., G. M. Converse & H. C. Dillon. 1980, "Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life", *J. Infect. Dis.,* vol. 142, no. 6, 923-933.

Greenberg, D., N. Givon-Lavi, A. Broides, I. Blancovich, N. Peled & R. Dagan. 2006, "The contribution of smoking and exposure to tobacco smoke to *Streptococcus pneumoniae* and *Haemophilus influenzae* carriage in children and their mothers", *Clin. Infect. Dis,* vol. 42, no. 7, 897-903.

Grennan, T., D. Leto, K. Luinstra, A. Petrich, M. Smieja, M. Loeb & J. Johnstone Prevalence of Streptococcus pneumoniae Colonization in Nursing Home Elderly: Comparison of Conventional Culture and Polymerase Chain Reaction. Infectious Diseases Society of America (IDSA) 48th Annual Meeting, Vancouver, Canada.

Griffith, F. 1928, "The significance of pneumococcal types", *J. Hyg.,* vol. 27, no. 2, 113-159.

Grimwood, K., V. A. Anderson, L. Bond, C. Catroppa, R. L. Hore, E. H. Keir, T. Nolan & D. M. Roberton. 1995, "Adverse outcomes of bacterial meningitis in school-age survivors", *Pediatrics,* vol. 95, no. 5, 646-656.

Guiral, S., T. J. Mitchell, B. Martin & J.-P. Claverys. 2005, "Competence-programmed predation of noncompetent cells in the human pathogen *Streptococcus pneumoniae*: genetic requirements.", *Proc. Nat. Acad. Sci. U. S. A.,* vol. 102, no. 24, 8710-8715.

Gundel, M. & G. Okura. 1933, "Untersuchungen über das gleichzeitige Vorkommen mehrerer Pneumokokkentypen bei Gesunden und ihre Bedeutung für die Epidemiologie", *Med. Microbiol. Immun.,* vol. 114, no. 4, 678-704.

Haas, W., C. K. Hesje, C. M. Sanfilippo & T. W. Morris. 2011, "High proportion of nontypeable *Streptococcus pneumoniae* isolates among sporadic, nonoutbreak cases of bacterial conjunctivitis", *Curr. Eye Res.*, vol. 36, no. 12, 1078-1085.

Hagelueken, G., H. Huang, I. Mainprize, C. Whitfield & J. Naismith. 2009, "Crystal structures of Wzb of *Escherichia coli* and CpsB of *Streptococcus pneumoniae,* representatives of two families of tyrosine phosphatases that regulate capsule assembly", *J. Mol. Biol.*, vol. 392, no. 3, 678-688.

Håkansson, A., H. Roche, S. Mirza, L. S. McDaniel, A. Brooks-Walter & D. E. Briles. 2001, "Characterization of binding of human lactoferrin to pneumococcal surface protein A", *Infect. Immun.,* vol. 69, no. 5, 3372-3381.

Hakenbeck, R., N. Balmelle, B. Weber, C. Gardès, W. Keck & A. de Saizieu. 2001, "Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 69, no. 4, 2477-2486.

Hakenbeck, R., A. König, I. Kern, M. Van Der Linden, W. Keck, D. Billot-Klein, R. Legrand, B. Schoot & L. Gutmann. 1998, "Acquisition of five high-Mr penicillin-binding protein variants during transfer of high-level beta-lactam resistance from *Streptococcus mitis* to *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 180, no. 7, 1831-1840.

Hall-Stoodley, L., F. Z. Hu, A. Gieseke, L. Nistico, D. Nguyen, J. Hayes, M. Forbes, D. P. Greenberg, B. Dice, A. Burrows, P. A. Wackym, P. Stoodley, J. C. Post, G. D. Ehrlich & J. E. Kerschner. 2006, "Direct detection of bacterial biofilms on the middle-ear mucosa of children with chronic otitis media", *JAMA,* vol. 296, no. 2, 202-211.

Hall-Stoodley, L., L. Nistico, K. Sambanthamoorthy, B. Dice, D. Nguyen, W. J. Mershon, C. Johnson, F. Z. Hu, P. Stoodley, G. D. Ehrlich & J. C. Post. 2008, "Characterization of biofilm matrix, degradation by DNase treatment and evidence of capsule downregulation in *Streptococcus pneumoniae* clinical isolates", *BMC Microbiology,* vol. 8, no. 173.

Hamel, M., C. Greene, T. Chiller, P. Ouma, C. Polyak, K. Otieno, J. Williamson, Y. Shi, D. Feikin & B. Marston. 2008, "Does cotrimoxazole prophylaxis for the prevention of HIV-associated opportunistic infections select for resistant pathogens in Kenyan adults?", *A. J. Trop. Med. Hyg.,* vol. 79, no. 3, 320-330.

Hammerschmidt, S. (2007). Pneumococcal Virulence Factors and Adhesion Proteins Targeting the Host. Molecular Biology of Streptococci. R. Hakenbeck and S. Chhatwal. London, Tayler & Francis Ltd.**:** 141-204.

Hammerschmidt, S., G. Bethe, P. H. Remane & G. S. Chhatwal. 1999, "Identification of pneumococcal surface protein A as a lactoferrin-binding protein of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 67, no. 4, 1683-1687.

Hammerschmidt, S., S. Talay, P. Brandtzaeg & G. Chhatwal. 1997, "SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component", *Mol. Microbiol.*, vol. 25, no. 6, 1113-1124.

Hammerschmidt, S., S. Wolff, A. Hocke, S. Rosseau, E. Müller & M. Rohde. 2005, "Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells.", *Infect. Immun.*, vol. 73, no. 8, 4653-4667.

Hammitt, L. L., D. L. Bruden, J. C. Butler, H. C. Baggett, D. A. Hurlburt, A. Reasonover & T. W. Hennessy. 2006, "Indirect effect of conjugate vaccine on adult carriage of *Streptococcus pneumoniae*: an explanation of trends in invasive pneumococcal disease.", *J. Infect. Dis.*, vol. 193, no. 11, 1487-1494.

Hampton, L. M., M. M. Farley, W. Schaffner, A. Thomas, A. Reingold, L. H. Harrison, R. Lynfield, N. M. Bennett, S. Petit, K. Gershman, J. Baumbach, B. Beall, J. Jorgensen, A. Glennen, E. R. Zell & M. Moore. 2011, "Prevention of antibiotic-nonsusceptible *Streptococcus pneumoniae* with conjugate vaccines.", *J. Infect. Dis.*, vol. 205, no. 3, 401-411.

Hanage, W., C. Bishop, S. Huang, A. Stevenson, S. Pelton, M. Lipsitch & J. Finkelstein. 2011, "Carried pneumococci in Massachusetts children: the contribution of clonal expansion and serotype switching", *Pediatr. Infect. Dis. J.*, vol. 30, no. 4, 302-.308

Hanage, W., J. Finkelstein, S. Huang, S. Pelton, A. Stevenson, K. Kleinman, V. Hinrichsen & C. Fraser. 2010, "Evidence that pneumococcal serotype replacement in Massachusetts following conjugate vaccination is now complete", *Epidemics*, vol. 2, no. 2, 80-84.

Hanage, W., C. Fraser, J. Tang, T. Connor & J. Corander. 2009, "Hyper-recombination, diversity, and antibiotic resistance in pneumococcus", *Science*, vol. 324, no. 5933, 1454-1457.

Hanage, W., S. Huang, M. Lipsitch, C. Bishop, D. Godoy, S. Pelton, R. Goldstein, H. Huot & J. Finkelstein. 2007, "Diversity and antibiotic resistance among nonvaccine serotypes of *Streptococcus pneumoniae* carriage isolates in the post-heptavalent conjugate vaccine era", *J. Infect. Dis.*, vol. 195, no. 3, 347-352.

Hanage, W. P., T. Kaijalainen, E. Herva, A. Saukkoriipi, R. Syrjänen & B. G. Spratt. 2005, "Using multilocus sequence data to define the pneumococcus", *J. Bacteriol.*, vol. 187, no. 17, 6223-6230.

Hanage, W. P., T. Kaijalainen, A. Saukkoriipi, J. L. Rickcord & B. G. Spratt. 2006, "A successful, diverse disease-associated lineage of nontypeable pneumococci that has lost the capsular biosynthesis locus", *J. Clin. Microbiol.*, vol. 44, no. 3, 743-749.

Hanquet, G., E. Kissling, A. Fenoll, R. George, A. Lepoutre, T. Lernout, D. Tarragó, E. Varon & J. Verhaegen. 2010, "Pneumococcal serotypes in children in 4 European countries", *Emerg. Infect. Dis.*, vol. 16, no. 9, 1428-1439.

Hansman, D. & M. Bullen. 1967, "A resistant pneumococcus", *Lancet*, vol. 2, no. 7295, 264-265.

Hansman, D., L. Devitt, H. Miles & I. Riley. 1974, "Pneumococci relatively insensitive to penicillin in Australia and New Guinea", *Med. J. Aust.*, vol. 2, no. 10, 353-356.

Hansman, D., S. Morris, M. Gregory & B. McDonald. 1985, "Pneumococcal carriage amongst Australian aborigines in Alice Springs, Northern Territory", *J. Hyg.*, vol. 95, no. 3, 677-684.

Hanson, B. R., B. A. Lowe & M. N. Neely. 2011, "Membrane topology and DNA-binding ability of the Streptococcal CpsA protein.", *J. Bacteriol.*, vol. 193, no. 2, 411-420.

Hardy, G., A. Magee, C. Ventura, M. Caimano & J. Yother. 2001, "Essential role for cellular phosphoglucomutase in virulence of type 3 *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 69, no. 4, 2309-2317.

Hare, K. M., P. Morris, H. Smith-Vaughan & A. J. Leach. 2008, "Random colony selection versus colony morphology for detection of multiple pneumococcal serotypes in nasopharyngeal swabs.", *Pediatr. Infect. Dis. J.*, vol. 27, no. 2, 178-180.

Hare, K. M., H. C. Smith-Vaughan & A. J. Leach. 2011, "Viability of respiratory pathogens cultured from nasopharyngeal swabs stored for up to 12years at −70°C in skim milk tryptone glucose glycerol broth", *J. Microbiol. Methods*, vol. no. 1-4.

Harf-Monteil, C., C. Granello, C. Le Brun, H. Monteil & P. Riegel. 2006, "Incidence and pathogenic effect of *Streptococcus pseudopneumoniae*", *J. Clin. Microbiol.*, vol. 44, no. 6, 2240-2241.

Harvey, R. M., A. D. Ogunniyi, A. Y. Chen & J. C. Paton. 2011, "Pneumolysin with low hemolytic activity confers an early growth advantage to *Streptococcus pneumoniae* in the blood", *Infect. Immun.*, vol. 79, no. 10, 4122-4130.

Hathaway, L., P. Bättig & K. Mühlemann. 2007, "In vitro expression of the first capsule gene of *Streptococcus pneumoniae*, *cpsA*, is associated with serotype-specific colonization prevalence and invasiveness", *Microbiology,* vol. 153, no. 8, 2465-2471.

Hathaway, L. J., P. Stutzmann Meier, P. Bättig, S. Aebi & K. Mühlemann. 2004, "A homologue of *aliB* is found in the capsule region of nonencapsulated *Streptococcus pneumoniae*", *J. Bacteriol.*, vol. 186, no. 12, 3721-3729.

Hausdorff, W. P. 2007, "The roles of pneumococcal serotypes 1 and 5 in paediatric invasive disease", *Vaccine,* vol. 25, no. 13, 2406-2412.

Hauser, C., S. Aebi & K. Mühlemann. 2004, "An internationally spread clone of *Streptococcus pneumoniae* evolves from low-level to higher-level penicillin resistance by uptake of penicillin-binding protein gene fragments from nonencapsulated pneumococci", *Antimicrob. Agents Chemother.,* vol. 48, no. 9, 3563-3566.

Hava, D. 2002, "Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors", *Mol. Microbiol.*, vol. 45, no. 5, 1389-1406.

Håvarstein, L., B. Martin, O. Johnsborg, C. Granadel & J. Claverys. 2006a, "New insights into the pneumococcal fratricide: relationship to clumping and identification of a novel immunity factor", *Mol. Microbiol.,* vol. 59, no. 4, 1297-1037.

Havarstein, L. S. (1999). Quorum Sensing and Peptide Pheromones in Streptococcal Competence for Genetic Transformation. Cell-Cell Signalling in Bacteria. G. M. Dunny and S. C. Winans. Washington, D.C., ASM Press**:** 9-26.

Håvarstein, L. S., G. Coomaraswamy & D. A. Morrison. 1995, "An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae*", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 92, no. 24, 11140-11144.

Håvarstein, L. S., D. B. Diep & I. F. Nes. 1995, "A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export", *Mol. Microbiol.*, vol. 16, no. 2, 229-240.

Håvarstein, L. S., P. Gaustad, I. F. Nes & D. A. Morrison. 1996, "Identification of the streptococcal competence-pheromone receptor", *Mol. Microbiol.*, vol. 21, no. 4, 863-869.

Håvarstein, L. S., R. Hakenbeck & P. Gaustad. 1997, "Natural competence in the genus *Streptococcus*: evidence that streptococci can change pherotype by interspecies recombinational exchanges", *J. Bacteriol.*, vol. 179, no. 21, 6589-6594.

Håvarstein, L. S., H. Holo & I. F. Nes. 1994, "The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria", *Microbiology*, vol. 140, no. 9, 2383-2389.

Håvarstein, L. S., B. Martin, O. Johnsborg, C. Granadel & J.-P. Claverys. 2006b, "New insights into the pneumococcal fratricide: relationship to clumping and identification of a novel immunity factor", *Mol. Microbiol.*, vol. 59, no. 4, 1297-1307.

Health Protection Agency. (2011). UK Standards for Microbiology Investigations: Bile Solubility Test.

Henrichsen, J. 1979, "The pneumococcal typing system and pneumococcal surveillance", *J. Infect.*, vol. 1, no. 31-37.

Henrichsen, J. 1995, "Six newly recognized types of *Streptococcus pneumoniae*", *J. Clin. Microbiol.*, vol. 33, no. 10, 2759-2762.

Henrichsen, J., E. Berntsson & B. Kaijser. 1980, "Comparison of counterimmunoelectrophoresis and the capsular reaction test for typing of pneumococci.", *J. Clin. Microbiol.*, vol. 11, no. 6, 589-592.

Hicks, L., L. Harrison, B. Flannery, J. Hadler, W. Schaffner, A. Craig, D. Jackson, A. Thomas, B. Beall & R. Lynfield. 2007, "Incidence of pneumococcal disease

due to non-pneumococcal conjugate vaccine (PCV7) serotypes in the United States during the era of widespread PCV7 vaccination, 1998–2004", *J. Infect. Dis.*, vol. 196, no. 9, 1346-1354.

Hill, P., Y. Cheung, A. Akisanya, K. Sankareh, G. Lahai, B. Greenwood & R. Adegbola. 2008, "Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study", *Clin. Infect. Dis.*, vol. 46, no. 6, 807-814.

Hill, P. C., A. Akisanya, K. Sankareh, Y.-B. Cheung, M. Saaka, G. Lahai, B. M. Greenwood & R. A. Adegbola. 2006, "Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian villagers.", *Clin. Infect. Dis.*, vol. 43, no. 6, 673-679.

Hiller, N., A. Ahmed, E. Powell, D. Martin, R. Eutsey, J. Earl, B. Janto, R. Boissy, J. Hogg & K. Barbadora. 2010, "Generation of genic diversity among *Streptococcus pneumoniae* strains via horizontal gene transfer during a chronic polyclonal pediatric infection", *PLoS Pathogens*, vol. 6, no. 9, e1001108.

Hiller, N., B. Janto, J. Hogg, R. Boissy, S. Yu, E. Powell, R. Keefe, N. Ehrlich, K. Shen & J. Hayes. 2007, "Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome", *J. Bacteriol.*, vol. 189, no. 22, 8186-8195.

Hiller, N. L., R. A. Eutsey, E. Powell, J. P. Earl, B. Janto, D. P. Martin, S. Dawid, A. Ahmed, M. J. Longwell, M. E. Dahlgren, S. Ezzo, H. Tettelin, S. C. Daugherty, T. J. Mitchell, T. A. Hillman, F. J. Buchinsky, A. Tomasz, H. de Lencastre, R. Sá-Leão, J. C. Post, F. Z. Hu & G. D. Ehrlich. 2011, "Differences in genotype and virulence among four multidrug-resistant *Streptococcus pneumoniae* isolates belonging to the PMEN1 clone.", *PLoS ONE*, vol. 6, no. 12, e28850.

Hoa, M., M. Syamal, L. Sachdeva, R. Berk & J. Coticchia. 2009a, "Demonstration of nasopharyngeal and middle ear mucosal biofilms in an animal model of acute otitis media.", *Ann. Otol. Rhinol. Laryngol.*, vol. 118, no. 4, 292-298.

Hoa, M., S. Tomovic, L. Nistico, L. Hall-Stoodley, P. Stoodley, L. Sachdeva, R. Berk & J. M. Coticchia. 2009b, "Identification of adenoid biofilms with middle ear pathogens in otitis-prone children utilizing SEM and FISH", *Int. J. Pediatr. Otorhinolaryngol.*, vol. 73, no. 9, 1242-1248.

Hogberg, L., P. Geli, H. Ringberg, E. Melander, M. Lipsitch & K. Ekdahl. 2007, "Age-and serogroup-related differences in observed durations of

nasopharyngeal carriage of penicillin-resistant pneumococci", *J. Clin. Microbiol.*, vol. 45, no. 3, 948-952.

Hoge, C. W., M. R. Reichler, E. A. Dominguez, J. C. Bremer, T. D. Mastro, K. A. Hendricks, D. M. Musher, J. A. Elliott, R. R. Facklam & R. F. Breiman. 1994, "An epidemic of pneumococcal disease in an overcrowded, inadequately ventilated jail.", *N. Engl. J. of Med.*, vol. 331, no. 10, 643-648.

Hogg, J., F. Hu, B. Janto, R. Boissy, J. Hayes, R. Keefe, J. Post & G. Ehrlich. 2007, "Characterization and modeling of the *Haemophilus influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains", *Genome Biol.*, vol. 8, no. 6, R103.

Holliday, M. G. 1981, "Pneumococcal typing by polyvalent counterimmunoelectrophoresis (PIE)", *J. Immun. Methods*, vol. 46, no. 2, 243-249.

Hollingshead, S., R. Becker & D. Briles. 2000, "Diversity of PspA: mosaic genes and evidence for past recombination in *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 68, no. 10, 5889-5900.

Holmberg, H., D. Danielsson, J. Hardie, A. Krook & R. Whiley. 1985, "Cross-reactions between alpha-streptococci and Omniserum, a polyvalent pneumococcal serum, demonstrated by direct immunofluorescence, immunoelectroosmophoresis, and latex agglutination", *J. Clin. Microbiol.*, vol. 21, no. 5, 745-748.

Holmes, A. R., R. McNab, K. W. Millsap, M. Rohde, S. Hammerschmidt, J. L. Mawdsley & H. F. Jenkinson. 2001, "The *pavA* gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence", *Mol. Microbiol.*, vol. 41, no. 6, 1395-1408.

Hoshino, T., T. Fujiwara & M. Kilian. 2005, "Use of phylogenetic and phenotypic analyses to identify nonhemolytic streptococci isolated from bacteremic patients", *J. Clin. Microbiol.*, vol. 43, no. 12, 6073-6085.

Hoskins, J., W. E. Alborn, J. Arnold, L. C. Blaszczak, S. Burgett, B. S. DeHoff, S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H. Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P. Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I. Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun, M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R. Jaskunas, P. R. Rosteck, P. L. Skatrud & J. I.

Glass. 2001, "Genome of the bacterium *Streptococcus pneumoniae* strain R6", *J. Bacteriol.,* vol. 183, no. 19, 5709-5717.

Houldsworth, S., P. W. Andrew & T. J. Mitchell. 1994, "Pneumolysin stimulates production of tumor necrosis factor alpha and interleukin-1 beta by human mononuclear phagocytes", *Infect. Immun.,* vol. 62, no. 4, 1501-1503.

Howe, J. G. & T. S. Wilson. 1972, "Co-trimoxazole-resistant pneumococci", *Lancet,* vol. 2, no. 7769, 184-185.

Hsieh, Y.-C., J.-T. Wang, W.-S. Lee, P.-R. Hsueh, P.-L. Shao, L.-Y. Chang, C.-Y. Lu, C.-Y. Lee, F.-Y. Huang & L.-M. Huang. 2006, "Serotype competence and penicillin resistance in *Streptococcus pneumoniae*", *Emerg. Infect. Dis.,* vol. 12, no. 11, 1709-1714.

Hsu, H. E., K. A. Shutt, M. R. Moore, B. W. Beall, N. M. Bennett, A. S. Craig, M. M. Farley, J. H. Jorgensen, C. A. Lexau, S. Petit, A. Reingold, W. Schaffner, A. Thomas, C. G. Whitney & L. H. Harrison. 2009, "Effect of pneumococcal conjugate vaccine on pneumococcal meningitis.", *N. Engl. J. Med.,* vol. 360, no. 3, 244-256.

Huang, S. S., V. L. Hinrichsen, A. E. Stevenson, S. L. Rifas-Shiman, K. Kleinman, S. I. Pelton, M. Lipsitch, W. P. Hanage, G. M. Lee & J. A. Finkelstein. 2009, "Continued impact of pneumococcal conjugate vaccine on carriage in young children", *Pediatrics,* vol. 124, no. 1, e1-11.

Huebner, R. E., R. Dagan, N. Porath, A. D. Wasas & K. P. Klugman. 2000a, "Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes", *Pediatr. Infect. Dis. J.,* vol. 19, no. 10, 1017-1020.

Huebner, R. E., A. D. Wasas, M. Hockman, K. P. Klugman & E. S. Group. 2003, "Bacterial aetiology of non-resolving otitis media in South African children.", *J. Laryngol. Otol.,* vol. 117, no. 3, 169-172.

Huebner, R. E., A. D. Wasas, A. Mushi, L. Mazhani, and K. Klugman. 1998, "Nasophayryngeal carriage and antimicrobial resistance in isolates of *Streptococcus pneumoniae* and *Haemophilus influenzae* type b in children under 5 years of age in Botswana", *Int. J. Infect. Dis.,* vol. 3, no. 1, 18-25.

Huebner, R. E., A. D. Wasas, K. P. Klugman & P. S. Group. 2000b, "Prevalence of nasopharyngeal antibiotic-resistant pneumococcal carriage in children

attending private paediatric practices in Johannesburg.", *S. Afr. Med. J.*, vol. 90, no. 11, 1116-1121.

Hui, F. M. & D. A. Morrison. 1991, "Genetic transformation in *Streptococcus pneumoniae*: nucleotide sequence analysis shows *comA*, a gene required for competence induction, to be a member of the bacterial ATP-dependent transport protein family", *J. Bacteriol.*, vol. 173, no. 1, 372-381.

Hui, F. M., L. Zhou & D. A. Morrison. 1995, "Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes", *Gene,* vol. 153, no. 1, 25-31.

Hussain, M., A. Melegaro, R. G. Pebody, R. George, W. J. Edmunds, R. Talukdar, S. A. Martin, A. Efstratiou & E. Miller. 2005, "A longitudinal household study of *Streptococcus pneumoniae* nasopharyngeal carriage in a UK setting", *Epidemiol. Infect.,* vol. 133, no. 5, 891-898.

Hyams, C., J. Yuste, K. Bax, E. Camberlein, J. Weiser & J. Brown. 2010, "*Streptococcus pneumoniae* resistance to complement-mediated immunity is dependent on the capsular serotype", *Infect. Immun.,* vol. 78, no. 2, 716-725.

Iannelli, F., M. R. Oggioni & G. Pozzi. 2002, "Allelic variation in the highly polymorphic locus *pspC* of *Streptococcus pneumoniae*", *Gene,* vol. 284, no. 1-2, 63-71.

Iannelli, F., M. R. Oggioni & G. Pozzi. 2005, "Sensor domain of histidine kinase ComD confers competence pherotype specificity in *Streptoccoccus pneumoniae*", *FEMS Microbiol. Lett.*, vol. 252, no. 2, 321-326.

Imperi, F., L. C. S. Antunes, J. Blom, L. Villa, M. Iacono, P. Visca & A. Carattoli. 2011, "The genomics of *Acinetobacter baumannii*: insights into genome plasticity, antimicrobial resistance and pathogenicity.", *IUBMB Life,* vol. 63, no. 12, 1068-1074.

Ing, J., E. O. Mason, S. L. Kaplan, L. B. Lamberth, P. A. Revell, R. A. Luna & K. G. Hulten. 2012, "Characterization of non-typeable and atypical *Streptococcus pneumoniae* pediatric isolates from 1994-2010", *J. Clin. Microbiol.,* vol. 50, no. 4, 1326-1330.

Inostroza, J., O. Trucco, V. Prado, A. M. Vinet, G. Retamal, G. Ossa, R. R. Facklam & R. U. Sorensen. 1998, "Capsular serotype and antibiotic resistance of *Streptococcus pneumoniae* isolates in two Chilean cities", *Clin. Diagn. Lab Immunol.*, vol. 5, no. 2, 176-180.

Inverarity, D., M. Diggle, R. Ure, P. Johnson, P. Altstadt, T. Mitchell, G. Edwards & D. Santana-Hernandez. 2011, "Molecular epidemiology and genetic diversity of pneumococcal carriage among children in Beni State, Bolivia", *Trans. R. Soc. Trop. Med. Hyg,* vol. 105, no. 8, 445-451.

Inverarity, D., M. Diggle, R. Ure, D. Santana-Hernandez, P. Altstadt, T. Mitchell & G. Edwards. 2010, "Simultaneous nasopharyngeal carriage of two pneumococcal multilocus sequence types with a serotype 3 phenotype", *Int. J. Microbiol.,* vol. 2010, no. 765479.

Iraurgui, P., M. J. Torres, A. Gandia, I. Vazquez, E. G. Cabrera, I. Obando, J. Garnacho & J. Aznar. 2010, "Modified sequential multiplex PCR for determining capsular serotypes of invasive pneumococci recovered from Seville", *Clin. Microbiol. Infect,* vol. 16, no. 9, 1504-1507.

Jackson, L. A., K. M. Neuzil, O. Yu, P. Benson, W. E. Barlow, A. L. Adams, C. A. Hanson, L. D. Mahoney, D. K. Shay, W. W. Thompson & V. S. Datalink. 2003, "Effectiveness of pneumococcal polysaccharide vaccine in older adults", *N. Engl. J. Med.,* vol. 348, no. 18, 1747-1755.

Jambo, K. C., E. Sepako, R. S. Heyderman & S. B. Gordon. 2010, "Potential role for mucosally active vaccines against pneumococcal pneumonia", *Trends Microbiol.,* vol. 18, no. 2, 81-89.

Jedrzejas, M. J. 2001, "Pneumococcal virulence factors: structure and function", *Microbiol. Mol. Biol. Rev.,* vol. 65, no. 2, 187-207.

Jedrzejas, M. J. 2004, "Extracellular virulence factors of *Streptococcus pneumoniae*", *Front. Biosci.,* vol. 9, no. 891-914.

Jefferies, J., L. Nieminen, L.-A. Kirkham, C. Johnston, A. Smith & T. J. Mitchell. 2007a, "Identification of a secreted cholesterol-dependent cytolysin (mitilysin) from *Streptococcus mitis*.", *J. Bacteriol.,* vol. 189, no. 2, 627-632.

Jefferies, J. M., E. Macdonald, S. N. Faust & S. C. Clarke. 2011, "13-valent pneumococcal conjugate vaccine (PCV13)", *Hum. Vaccin.,* vol. 7, no. 10, 1012-1018.

Jefferies, J. M., A. J. Smith, G. F. S. Edwards, J. Mcmenamin, T. J. Mitchell & S. C. Clarke. 2010a, "Temporal analysis of invasive pneumococcal clones from Scotland illustrates fluctuations in diversity of serotype and genotype in the absence of pneumococcal conjugate vaccine", *J. Clin. Microbiol.,* vol. 48, no. 1, 87-96.

Jefferies, J. M. C., C. H. G. Johnston, L.-A. S. Kirkham, G. J. M. Cowan, K. S. Ross, A. Smith, S. C. Clarke, A. B. Brueggemann, R. C. George, B. Pichon, G. Pluschke, V. Pfluger & T. J. Mitchell. 2007b, "Presence of nonhemolytic pneumolysin in serotypes of *Streptococcus pneumoniae* associated with disease outbreaks.", *J. Infect. Dis.,* vol. 196, no. 6, 936-944.

Jefferies, J. M. C., A. Smith, S. C. Clarke, C. Dowson & T. J. Mitchell. 2004, "Genetic analysis of diverse disease-causing pneumococci indicates high levels of diversity within serotypes and capsule switching.", *J. Clin. Microbiol.,* vol. 42, no. 12, 5681-5688.

Jefferies, J. M. C., W. S. N. Tee & S. C. Clarke. 2011, "Molecular analysis of *Streptococcus pneumoniae* clones causing invasive disease in children in Singapore.", *J. Med. Microbiol.,* vol. 60, no. 6, 750-755.

Jefferies, J. M. C., A. S. Tocheva, H. Rubery, J. Bennett, J. Garland, M. Christodoulides, S. N. Faust, A. Smith, T. J. Mitchell & S. C. Clarke. 2010b, "Identification of novel pneumolysin alleles from paediatric carriage isolates of *Streptococcus pneumoniae*.", *J. Med. Microbiol.,* vol. 59, no. 7, 808-814.

Jiang, S. M., L. Wang & P. R. Reeves. 2001, "Molecular characterization of *Streptococcus pneumoniae* type 4, 6B, 8, and 18C capsular polysaccharide gene clusters", *Infect. Immun.,* vol. 69, no. 3, 1244-1255.

Johnsborg, O., V. Eldholm, M. L. Bjørnstad & L. S. Håvarstein. 2008, "A predatory mechanism dramatically increases the efficiency of lateral gene transfer in *Streptococcus pneumoniae* and related commensal species", *Mol. Microbiol.,* vol. 69, no. 1, 245-253.

Johnsborg, O. & L. Håvarstein. 2009, "Regulation of natural genetic transformation and acquisition of transforming DNA in *Streptococcus pneumoniae*", *FEMS Microbiol. Rev.,* vol. 33, no. 3, 627-642.

Johnsborg, O., P. E. Kristiansen, T. Blomqvist & L. S. Håvarstein. 2006, "A hydrophobic patch in the competence-stimulating peptide, a pneumococcal competence pheromone, is essential for specificity and biological activity", *J. Bacteriol.,* vol. 188, no. 5, 1744-1749.

Johnson, H. L., M. Deloria-Knoll, O. S. Levine, S. K. Stoszek, L. Freimanis Hance, R. Reithinger, L. R. Muenz & K. L. O'Brien. 2010, "Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project.", *PLoS Med.,* vol. 7, no. 10. e1000348.

Johnston, C., J. Hinds, A. Smith, M. van der Linden, J. Van Eldere & T. J. Mitchell. 2010, "Detection of large numbers of pneumococcal virulence genes in streptococci of the mitis group", *J. Clin. Microbiol.,* vol. 48, no. 8, 2762-2769.

Johnston, J. W., L. E. Myers, M. M. Ochs, W. H. Benjamin, D. E. Briles & S. K. Hollingshead. 2004, "Lipoprotein PsaA in virulence of *Streptococcus pneumoniae*: surface accessibility and role in protection from superoxide", *Infect. Immun.,* vol. 72, no. 10, 5858-5867.

Joloba, M. L., S. Bajaksouzian, E. Palavecino, C. Whalen & M. R. Jacobs. 2001, "High prevalence of carriage of antibiotic-resistant *Streptococcus pneumoniae* in children in Kampala, Uganda", *Int. J. Antimicrob. Agents,* vol. 17, no. 5, 395-400.

Jourdain, S., P. R. Smeesters, O. Denis, M. Dramaix, V. Sputael, X. Malaviolle, L. Van Melderen & A. Vergison. 2011, "Differences in nasopharyngeal bacterial carriage in preschool children from different socio-economic origins.", *Clin. Microbiol. Infect,* vol. 17, no. 6, 907-914.

Jousimies-Somer, H. R., S. Savolainen & J. S. Ylikoski. 1989, "Comparison of the nasal bacterial floras in two groups of healthy subjects and in patients with acute maxillary sinusitis.", *J. Clin. Microbiol.*, vol. 27, no. 12, 2736-2743.

Kadioglu, A., H. Brewin, T. Härtel, J. L. Brittan, M. Klein, S. Hammerschmidt & H. F. Jenkinson. 2010, "Pneumococcal protein PavA is important for nasopharyngeal carriage and development of sepsis", *Mol. Oral Microbiol.,* vol. 25, no. 1, 50-60.

Kadioglu, A., A. M. Cuppone, C. Trappetti, T. List, A. Spreafico, G. Pozzi, P. W. Andrew & M. R. Oggioni. 2011, "Sex-based differences in susceptibility to respiratory and systemic pneumococcal disease in mice", *J. Infect. Dis.*, vol. 204, no. 12, 1971-1979.

Kadioglu, A., S. Taylor, F. Iannelli, G. Pozzi, T. J. Mitchell & P. W. Andrew. 2002, "Upper and lower respiratory tract infection by *Streptococcus pneumoniae* is affected by pneumolysin deficiency and differences in capsule type", *Infect. Immun.,* vol. 70, no. 6, 2886-2890.

Kadioglu, A., J. N. Weiser, J. C. Paton & P. W. Andrew. 2008, "The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease", *Nat. Rev. Microbiol.,* vol. 6, no. 4, 288-301.

Kaijalainen, T., E. Ruokokoski, P. Ukkonen & E. Herva. 2004, "Survival of *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Moraxella catarrhalis* frozen in skim milk-tryptone-glucose-glycerol medium", *J. Clin. Microbiol.*, vol. 42, no. 1, 412-414.

Kaldor, J., R. Asznowicz & B. Dwyer. 1988, "Serotyping of *Streptococcus pneumoniae* by latex agglutination.", *Pathology*, vol. 20, no. 1, 45-47.

Kaltoft, M. S., U. B. S. Sørensen, H.-C. Slotved & H. Bossen Konradsen. 2008, "An easy method for detection of nasopharyngeal carriage of multiple *Streptococcus pneumoniae* serotypes", *J. Microbiol. Methods*, vol. 75, no. 3, 540-544.

Kamerling, J. P. (2000). Pneumococcal Polysaccharides: A Chemical View. *Streptococcus pneumoniae*: Molecular Biology & Mechanisms of Disease. A. Tomasz. New York, Mary Ann Liebert, Inc.**:** 81-114.

Karlowsky, J., P. Lagacé-Wiens, D. Low & G. Zhanel. 2009, "Annual macrolide prescription rates and the emergence of macrolide resistance among *Streptococcus pneumoniae* in Canada from 1995 to 2005", *Int. J. Antimicrob. Agents*, vol. 34, no. 4, 375-379.

Kausmally, L., O. Johnsborg, M. Lunde, E. Knutsen & L. S. Håvarstein. 2005, "Choline-binding protein D (CbpD) in *Streptococcus pneumoniae* is essential for competence-induced cell lysis", *J. Bacteriol.*, vol. 187, no. 13, 4338-4345.

Kawamura, Y., X. G. Hou, F. Sultana, H. Miura & T. Ezaki. 1995, "Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus*", *Int. J. Syst. Bacteriol.*, vol. 45, no. 2, 406-408.

Kawamura, Y., R. A. Whiley, S. E. Shu, T. Ezaki & J. M. Hardie. 1999, "Genetic approaches to the identification of the mitis group within the genus *Streptococcus*", *Microbiology*, vol. 145, no. 9, 2605-2613.

Kaye, P., R. Malkani, S. Martin, M. Slack, C. Trotter, M. Jit, R. George & E. Miller (2009). Invasive Pneumococcal Disease (IPD) in England & Wales after 7-valent conjugate vaccine (PCV7); potential impact of 10 and 13-valent vaccines. The 27th Annual Meeting of the European Society for Paediatric Infectious Diseases. Brussels.

Keith, E. R. & D. R. Murdoch. 2008, "Antimicrobial susceptibility profile of *Streptococcus pseudopneumoniae* isolated from sputum", *Antimicrob. Agents Chemother.,* vol. 52, no. 8, 2998.

Keith, E. R., R. G. Podmore, T. P. Anderson & D. R. Murdoch. 2006, "Characteristics of *Streptococcus pseudopneumoniae* isolated from purulent sputum samples", *J. Clin. Microbiol.,* vol. 44, no. 3, 923-927.

Kellner, J. D., D. Scheifele, O. G. Vanderkooi, J. MacDonald, D. L. Church & G. J. Tyrrell. 2008, "Effects of routine infant vaccination with the 7-valent pneumococcal conjugate vaccine on nasopharyngeal colonization with *Streptococcus pneumoniae* in children in Calgary, Canada", *Pediatr. Infect. Dis. J.*, vol. 27, no. 6, 526-532.

Kelly, T., J. P. Dillard & J. Yother. 1994, "Effect of genetic switching of capsular type on virulence of *Streptococcus pneumoniae*", *Inject. Immun.,* vol. 62, no. 5, 1813-1819.

Kharat, A. & A. Tomasz. 2006, "Drastic reduction in the virulence of *Streptococcus pneumoniae* expressing type 2 capsular polysaccharide but lacking choline residues in the cell wall", *Mol. Microbiol.,* vol. 60, no. 1, 93-107.

Kikuchi, K., T. Enari, K. Totsuka & K. Shimizu. 1995, "Comparison of phenotypic characteristics, DNA-DNA hybridization results, and results with a commercial rapid biochemical and enzymatic reaction system for identification of viridans group streptococci.", *J. Clin. Microbiol.,* vol. 33, no. 5, 1215-1222.

Kilian, M. & K. Holmgren. 1981, "Ecology and nature of immunoglobulin A1 protease-producing streptococci in the human oral cavity and pharynx", *Infect. Immun.,* vol. 31, no. 3, 868-873.

Kilian, M., K. Poulsen, T. Blomqvist, L. S. Håvarstein, M. Bek-Thomsen, H. Tettelin & U. B. S. Sørensen. 2008, "Evolution of *Streptococcus pneumoniae* and its close commensal relatives", *PLoS ONE,* vol. 3, no. 7, e2683.

Kim, H. S., S.-J. Lee, H. J. Yoon, D. R. An, D. J. Kim, S.-J. Kim & S. W. Suh. 2011, "Crystal structures of YwqE from *Bacillus subtilis* and CpsB from *Streptococcus pneumoniae*, unique metal-dependent tyrosine phosphatases.", *J. Struct. Biol.,* vol. 175, no. 3, 442-450.

Kim, J. O., S. Romero-Steiner, U. B. Sørensen, J. Blom, M. Carvalho, S. Barnard, G. Carlone & J. N. Weiser. 1999, "Relationship between cell surface

carbohydrates and intrastrain variation on opsonophagocytosis of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 67, no. 5, 2327-2333.

Kim, J. O. & J. N. Weiser. 1998, "Association of intrastrain phase variation in quantity of capsular polysaccharide and teichoic acid with the virulence of *Streptococcus pneumoniae*", *J. Infect. Dis.*, vol. 177, no. 2, 368-377.

King, S., K. Hippe, J. Gould, D. Bae, S. Peterson, R. Cline, C. Fasching, E. Janoff & J. Weiser. 2004, "Phase variable desialylation of host proteins that bind to *Streptococcus pneumoniae* in vivo and protect the airway", *Mol. Microbiol.*, vol. 54, no. 1, 159-171.

King, S., K. Hippe & J. Weiser. 2006, "Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*", *Mol. Microbiol.*, vol. 59, no. 3, 961-974.

King, S. J., A. M. Whatmore & C. G. Dowson. 2005, "NanA, a neuraminidase from *Streptococcus pneumoniae*, shows high levels of sequence diversity, at least in part through recombination with *Streptococcus oralis*", *J. Bacteriol.*, vol. 187, no. 15, 5376-5386.

Kirkham, L.-A. S., J. M. C. Jefferies, A. R. Kerr, Y. Jing, S. C. Clarke, A. Smith & T. J. Mitchell. 2006, "Identification of invasive serotype 1 pneumococcal isolates that express nonhemolytic pneumolysin", *J. Clin. Microbiol.*, vol. 44, no. 1, 151-159.

Kirkman, J. B., Jr., J. Fischer & J. S. Pagano. 1970, "A microtiter plate technique for the agglutination typing of *Diplococcus pneumoniae*", *J. Infect. Dis.*, vol. 121, no. 2, 217-221.

Klugman, K. P., S. A. Madhi, R. E. Huebner, R. Kohberger, N. Mbelle, N. Pierce & V. T. Group. 2003, "A trial of a 9-valent pneumococcal conjugate vaccine in children with and those without HIV infection", *N. Engl. J. Med.*, vol. 349, no. 14, 1341-1348.

Kolkman, M. A., B. A. van der Zeijst & P. J. Nuijten. 1998, "Diversity of capsular polysaccharide synthesis gene clusters in *Streptococcus pneumoniae*", *J. Biochem.*, vol. 123, no. 5, 937-945.

Kong, F., M. Brown, A. Sabananthan, X. Zeng & G. L. Gilbert. 2006, "Multiplex PCR-based reverse line blot hybridization assay to identify 23 *Streptococcus pneumoniae* polysaccharide vaccine serotypes", *J. Clin. Microbiol.*, vol. 44, no. 5, 1887-1891.

Kong, F. & G. L. Gilbert. 2003, "Using *cpsA-cpsB* sequence polymorphisms and serotype-/group-specific PCR to predict 51 *Streptococcus pneumoniae* capsular serotypes", *J. Med. Microbiol.,* vol. 52, no. 12, 1047-1058.

Kong, F. & G. L. Gilbert. 2006, "Multiplex PCR-based reverse line blot hybridization assay (mPCR/RLB)--a practical epidemiological and diagnostic tool", *Nature Protoc.,* vol. 1, no. 6, 2668-2680.

Kong, F., L. Ma & G. L. Gilbert. 2005, "Simultaneous detection and serotype identification of *Streptococcus agalactiae* using multiplex PCR and reverse line blot hybridization.", *J. Med. Microbiol.,* vol. 54, no. 12, 1133-1138.

Kong, F., W. Wang, J. Tao, L. Wang, Q. Wang, A. Sabananthan & G. L. Gilbert. 2005, "A molecular-capsular-type prediction system for 90 *Streptococcus pneumoniae* serotypes using partial *cpsA-cpsB* sequencing and *wzy*- or *wzx*-specific PCR", *J. Med. Microbiol.,* vol. 54, no. 4, 351-356.

Kostyukova, N. N., M. O. Volkova, V. V. Ivanova & A. S. Kvetnaya. 1995, "A study of pathogenic factors of *Streptococcus pneumoniae* strains causing meningitis.", *FEMS Immun. Med. Microbiol.,* vol. 10, no. 2, 133-137.

Kowalko, J. E. & M. E. Sebert. 2008, "The *Streptococcus pneumoniae* competence regulatory system influences respiratory tract colonization", *Infect. Immun.,* vol. 76, no. 7, 3131-3140.

Kronvall, G. 1973, "A rapid slide-agglutination method for typing pneumococci by means of specific antibody adsorbed to protein A-containing staphylococci", *J. Med. Microbiol.,* vol. 6, no. 2, 187-190.

Kwambana, B. A., M. R. Barer, C. Bottomley, R. A. Adegbola & M. Antonio. 2011, "Early acquisition and high nasopharyngeal co-colonisation by *Streptococcus pneumoniae* and three respiratory pathogens amongst Gambian new-borns and infants", *BMC Infect. Dis.,* vol. 11, no. 175.

Lafong, A. C. & E. Crothers. 1988, "Simple latex agglutination method for typing pneumococci.", *J. Clin. Path.,* vol. 41, no. 2, 230-231.

Laible, G., B. G. Spratt & R. Hakenbeck. 1991, "Interspecies recombinational events during the evolution of altered PBP 2x genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*", *Mol. Microbiol.,* vol. 5, no. 8, 1993-2002.

Lalitha, M. K., R. Pai, T. J. John, K. Thomas, M. V. Jesudason, K. N. Brahmadathan, G. Sridharan & M. C. Steinhoff. 1996, "Serotyping of *Streptococcus pneumoniae* by agglutination assays: a cost-effective technique for developing countries", *Bull. World Health Organ.*, vol. 74, no. 4, 387-390.

Lalitha, M. K., K. Thomas, R. S. Kumar & M. C. Steinhoff. 1999, "Serotyping of Streptococcus pneumoniae by coagglutination with 12 pooled antisera", *Journal of Clinical Microbiology*, vol. 37, no. 1, 263-265.

Lau, G. W., S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison & D. W. Holden. 2001, "A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence", *Mol. Microbiol.*, vol. 40, no. 3, 555-571.

Laval, C. B., A. L. S. S. de Andrade, F. C. Pimenta, J. G. de Andrade, R. M. de Oliveira, S. A. Silva, E. C. de Lima, J. L. Fabio, S. T. Casagrande & M. C. C. Brandileone. 2006, "Serotypes of carriage and invasive isolates of *Streptococcus pneumoniae* in Brazilian children in the era of pneumococcal vaccines", *Clin. Microbiol. Infect.*, vol. 12, no. 1, 50-55.

Lawrence, E. R., C. A. Arias, B. Duke, D. Beste, K. Broughton, A. Efstratiou, R. C. George & L. M. Hall. 2000, "Evaluation of serotype prediction by *cpsA-cpsB* gene polymorphism in *Streptococcus pneumoniae*", *J. Clin. Microbiol.*, vol. 38, no. 4, 1319-1323.

Lawrence, E. R., D. B. Griffiths, S. A. Martin, R. C. George & L. M. C. Hall. 2003, "Evaluation of semiautomated multiplex PCR assay for determination of *Streptococcus pneumoniae* serotypes and serogroups", *J. Clin. Microbiol.*, vol. 41, no. 2, 601-607.

Leach, A. J. 1999, "Otitis media in Australian Aboriginal children: an overview", *Inte. J. Pediatr. Otorhinolaryngol.*, vol. 49 Suppl 1, no. S173-178.

Leach, A. J., J. B. Boswell, V. Asche, T. G. Nienhuys & J. D. Mathews. 1994, "Bacterial colonization of the nasopharynx predicts very early onset and persistence of otitis media in Australian aboriginal infants.", *Pediatr. Infect. Dis. J.*, vol. 13, no. 11, 983-989.

Leach, A. J., T. M. Shelby-James, M. Mayo, M. Gratten, A. C. Laming, B. J. Currie & J. D. Mathews. 1997, "A prospective study of the impact of community-based azithromycin treatment of trachoma on carriage and resistance of *Streptococcus pneumoniae*", *Clin. Infect. Dis.*, vol. 24, no. 3, 356-362.

Lee, C. J., K. Koizumi, J. Henrichsen, B. Perch, C. S. Lin & W. Egan. 1984, "Capsular polysaccharides of nongroupable streptococci that cross-react with pneumococcal group 19", *J. Immun.*, vol. 133, no. 5, 2706-2711.

Lee, J., J. Song & K. Ko. 2010, "Recombination rates of *Streptococcus pneumoniae* isolates with both *erm (B)* and *mef (A)* genes", *FEMS Microbiol. Lett.*, vol. 309, no. 2, 163-169.

Lee, M. S. & D. A. Morrison. 1999, "Identification of a new regulator in *Streptococcus pneumoniae* linking quorum sensing to competence for genetic transformation", *J. Bacteriol.*, vol. 181, no. 16, 5004-5016.

Leegaard, T. M., H. J. Bootsma, D. A. Caugant, M. J. Eleveld, T. Mannsåker, L. O. Frøholm, P. Gaustad, E. A. Høiby & P. W. M. Hermans. 2010, "Phenotypic and genomic characterization of pneumococcus-like streptococci isolated from HIV-seropositive patients", *Microbiology*, vol. 156, no. 3, 838-848.

Lefébure, T. & M. Stanhope. 2007, "Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition", *Genome Biol.*, vol. 8, no. 5, R71.

Leimkugel, J., A. Adams Forgor, S. Gagneux, V. Pflüger, C. Flierl, E. Awine, M. Naegeli, J.-P. Dangy, T. Smith, A. Hodgson & G. Pluschke. 2005, "An outbreak of serotype 1 *Streptococcus pneumoniae* meningitis in northern Ghana with features that are characteristic of *Neisseria meningitidis* meningitis epidemics", *J. Infect. Dis.*, vol. 192, no. 2, 192-199.

Leung, M., C. L. Ling, H. Ciesielczuk, J. Lockwood, S. Thurston, B. M. Charalambous & S. H. Gillespie. 2012, "*Streptococcus pseudopneumoniae* identification by pherotype: a method to assist the understanding of a potentially emerging or overlooked pathogen.", *J. Clin. Microbiol.*, vol. 50, no. 5, 1694-1690.

Leung, M. H. Y., N. M. Oriyo, S. H. Gillespie & B. M. Charalambous. 2011, "The adaptive potential during nasopharyngeal colonisation of *Streptococcus pneumoniae*", *Infect. Genet. and Evol.*, vol. 11, no. 8, 1989-1995.

Lexau, C. A., R. Lynfield, R. Danila, T. Pilishvili, R. Facklam, M. M. Farley, L. H. Harrison, W. Schaffner, A. Reingold, N. M. Bennett, J. Hadler, P. R. Cieslak, C. G. Whitney & A. B. C. S. Team. 2005, "Changing epidemiology of invasive pneumococcal disease among older adults in the era of pediatric pneumococcal conjugate vaccine", *JAMA*, vol. 294, no. 16, 2043-2051.

Li, Y.-H., N. Tang, M. B. Aspiras, P. C. Y. Lau, J. H. Lee, R. P. Ellen & D. G. Cvitkovitch. 2002, "A quorum-sensing signaling system essential for genetic competence in *Streptococcus mutans* is involved in biofilm formation", *J. Bacteriol.*, vol. 184, no. 10, 2699-2708.

Li, Y. H., P. C. Lau, J. H. Lee, R. P. Ellen & D. G. Cvitkovitch. 2001, "Natural genetic transformation of *Streptococcus mutans* growing in biofilms", *J. Bacteriol.*, vol. 183, no. 3, 897-908.

Lin, J., M. S. Kaltoft, A. P. Brandao, G. Echaniz-Aviles, M. C. C. Brandileone, S. K. Hollingshead, W. H. Benjamin & M. H. Nahm. 2006, "Validation of a multiplex pneumococcal serotyping assay with clinical samples", *J. Clin. Microbiol.*, vol. 44, no. 2, 383-388.

Liñares, J., T. Alonso, J. L. Pérez, J. Ayats, M. A. Domínguez, R. Pallarés & R. Martín. 1992, "Decreased susceptibility of penicillin-resistant pneumococci to twenty-four beta-lactam antibiotics", *J. Antimicrob. Chemother*, vol. 30, no. 3, 279-288.

Lipsitch, M., J. K. Dykes, S. E. Johnson, E. W. Ades, J. King, D. E. Briles & G. M. Carlone. 2000, "Competition among *Streptococcus pneumoniae* for intranasal colonization in a mouse model", *Vaccine*, vol. 18, no. 25, 2895-2901.

Llull, D., E. García & R. López. 2001, "Tts, a processive β-glucosyltransferase of *Streptococcus pneumoniae*, directs the synthesis of the branched type 37 capsular polysaccharide in pneumococcus and other Gram-positive species", *J. Biol. Chem.*, vol. 276, no. 24, 21053-21061.

Llull, D., R. López, E. García & R. Muñoz. 1998, "Molecular structure of the gene cluster responsible for the synthesis of the polysaccharide capsule of *Streptococcus pneumoniae* type 33F", *Biochim. Biophys. Acta.*, vol. 1443, no. 1-2, 217-224.

Llull, D., R. Muñoz, R. López & E. García. 1999, "A single gene (*tts*) located outside the cap locus directs the formation of *Streptococcus pneumoniae* type 37 capsular polysaccharide. Type 37 pneumococci are natural, genetically binary strains.", *J. Exp. Med.*, vol. 190, no. 2, 241-251.

Lo, W. T., C. C. Wang, C. M. Yu & M. L. Chu. 2003, "Rate of nasopharyngeal carriage, antimicrobial resistance and serotype of *Streptococcus pneumoniae* among children in northern Taiwan", *J. Microbiol. Immunol. Infect.*, vol. 36, no. 3, 175-181.

Lock, R. A., Q. Y. Zhang, A. M. Berry & J. C. Paton. 1996, "Sequence variation in the *Streptococcus pneumoniae* pneumolysin gene affecting haemolytic activity and electrophoretic mobility of the toxin.", *Microb. Pathog.*, vol. 21, no. 2, 71-83.

Lomholt, H. 1995, "Evidence of recombination and an antigenically diverse immunoglobulin A1 protease among strains of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 63, no. 11, 4238-4243.

Long, S. S. 2005, "Capsules, clones, and curious events: pneumococcus under fire from polysaccharide conjugate vaccine", *Clin. Infect. Dis.*, vol. 41, no. 1, 30-34.

Loo, C. Y., D. A. Corliss & N. Ganeshkumar. 2000, "*Streptococcus gordonii* biofilm formation: identification of genes that code for biofilm phenotypes", *J. Bacteriol.*, vol. 182, no. 5, 1374-1382.

López, R. 2006, "Pneumococcus: the sugar-coated bacteria", *Int. Microbiol.*, vol. 9, no. 3, 179-190.

López, R., C. Ronda & E. Garcia. 1990, "Autolysins are direct involved in the bactericidal effect caused by penicillin in wild type and in tolerant pneumococci.", *FEMS Microbiol. Lett.*, vol. 54, no. 1-3, 317-322.

Lu, Y., J. Gross, D. Bogaert, A. Finn, L. Bagrade, Q. Zhang, J. Kolls, A. Srivastava, A. Lundgren & S. Forte. 2008, "Interleukin-17A mediates acquired immunity to pneumococcal colonization", *PLoS Pathog.*, vol. 4, no. 9, e1000159.

Luo, P., H. Li & D. A. Morrison. 2003, "ComX is a unique link between multiple quorum sensing outputs and competence in *Streptococcus pneumoniae*", *Mol. Microbiol.*, vol. 50, no. 2, 623-633.

Luo, P. & D. A. Morrison. 2003, "Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in *Streptococcus pneumoniae*", *J. Bacteriol.*, vol. 185, no. 1, 349-358.

Lysenko, E. S., T. B. Clarke, M. Shchepetov, A. J. Ratner, D. I. Roper, C. G. Dowson & J. N. Weiser. 2007, "Nod1 signaling overcomes resistance of *S. pneumoniae* to opsonophagocytic killing.", *PLoS Pathog.*, vol. 3, no. 8, e118.

Lysenko, E. S., A. J. Ratner, A. L. Nelson & J. N. Weiser. 2005, "The role of innate immune responses in the outcome of interspecies competition for colonization of mucosal surfaces", *PLoS Pathog.,* vol. 1, no. 1, e1.

Mackenzie, G., J. Carapetis, A. J. Leach, K. Hare & P. Morris. 2007, "Indirect effects of childhood pneumococcal vaccination on pneumococcal carriage among adults and older children in Australian Aboriginal communities", *Vaccine,* vol. 25, no. 13, 2428-2433.

Mackenzie, G. A., A. J. Leach, J. R. Carapetis, J. Fisher & P. S. Morris. 2010, "Epidemiology of nasopharyngeal carriage of respiratory bacterial pathogens in children and adults: cross-sectional surveys in a population with high rates of pneumococcal disease.", *BMC Infect. Dis.,* vol. 10, no. 304.

Magee, A. D. & J. Yother. 2001, "Requirement for capsule in colonization by *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 69, no. 6, 3755-3761.

Maiden, M. C., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman & B. G. Spratt. 1998, "Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms", *Proc. Nat. Acad. Sci. U. S. A,* vol. 95, no. 6, 3140-3145.

Malley, R., A. Srivastava & M. Lipsitch. 2006, "Antibody-independent, interleukin-17A-mediated, cross-serotype immunity to pneumococci in mice immunized intranasally with the cell wall polysaccharide", *Infect. Immun.* , vol. 74. no. 4, 2187-2195.

Malley, R., K. Trzcinski, A. Srivastava, C. M. Thompson, P. W. Anderson & M. Lipsitch. 2005, "CD4+ T cells mediate antibody-independent acquired immunity to pneumococcal colonization", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 102, no. 13, 4848-4853.

Manco, S., F. Hernon, H. Yesilkaya, J. C. Paton, P. W. Andrew & A. Kadioglu. 2006, "Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis", *Infect. Immun.,* vol. 74, no. 7, 4014-4020.

Marra, A., & D. Brigham. 2001, "*Streptococcus pneumoniae* causes experimental meningitis following intranasal and otitis media infections via a nonhematogenous route", *Infect. Immun.*, vol. 69, no. 12, 7318-7325.

Margolis, E., A. Yates & B. R. Levin. 2010, "The ecology of nasal colonization of *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus*: the role of competition and interactions with host's immune response", *BMC Microbiol.,* vol. 10, no. 59.

Marimon, J. M., M. Ercibengoa, M. Alonso, M. Zubizarreta & E. Pérez-Trallero. 2009, "Clonal structure and 21-year evolution of *Streptococcus pneumoniae* serotype 1 isolates in northern Spain", *Clin. Microbiol. Infect,* vol. 15, no. 9, 875-877.

Marion, C., J. M. Stewart, M. F. Tazi, A. M. Burnaugh, C. M. Linke, S. A. Woodiga & S. J. King. 2012, "*Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth.", *Infect. Immun.,* vol. 80, no. 4, 1390-1398.

Marra, A. & D. Brigham. 2001, "*Streptococcus pneumoniae* causes experimental meningitis following intranasal and otitis media infections via a nonhematogenous route", *Infect. Immun.,* vol. 69, no. 12, 7318-7325.

Marra, A., S. Lawson, J. S. Asundi, D. Brigham & A. E. Hromockyj. 2002, "*In vivo* characterization of the *psa* genes from *Streptococcus pneumoniae* in multiple models of infection", *Microbiology,* vol. 148, no. 5, 1483-1491.

Marsh, R., H. Smith-Vaughan, K. M. Hare, M. Binks, F. Kong, J. Warning, G. L. Gilbert, P. Morris & A. J. Leach. 2010, "The nonserotypeable pneumococcus: phenotypic dynamics in the era of anticapsular vaccines", *J. Clin. Microbiol.,* vol. 48, no. 3, 831-835.

Martin, C., C. Sibold & R. Hakenbeck. 1992, "Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant *Streptococcus pneumoniae* isolated in South Africa and Spain", *EMBO J.,* vol. 11, no. 11, 3831-3836.

Martin, M., J. H. Turco, M. E. Zegans, R. R. Facklam, S. Sodha, J. A. Elliott, J. H. Pryor, B. Beall, D. D. Erdman, Y. Y. Baumgartner, P. A. Sanchez, J. D. Schwartzman, J. Montero, A. Schuchat & C. G. Whitney. 2003, "An outbreak of conjunctivitis due to atypical *Streptococcus pneumoniae*", *N. Engl. J. Med.,* vol. 348, no. 12, 1112-1121.

Martner, A., C. Dahlgren, J. C. Paton & A. E. Wold. 2008, "Pneumolysin released during *Streptococcus pneumoniae* autolysis is a potent activator of intracellular oxygen radical production in neutrophils", *Infect. Immun.,* vol. 76, no. 9, 4079-4087.

Marton, A., M. Gulyas, R. Munoz & A. Tomasz. 1991, "Extremely high incidence of antibiotic resistance in clinical isolates of *Streptococcus pneumoniae* in Hungary", *J. Infect. Dis.*, vol. 163, no. 3, 542-548.

Mascher, T., D. Zähner, M. Merai, N. Balmelle, A. B. de Saizieu & R. Hakenbeck. 2003, "The *Streptococcus pneumoniae cia* regulon: CiaR target sites and transcription profile analysis", *J. Bacteriol.*, vol. 185, no. 1, 60-70.

Masters, P. L., W. Brumfitt, R. L. Mendez & M. Likar. 1958, "Bacterial flora of the upper respiratory tract in Paddington families. 1952-4", *Br. Med. J.*, vol. 1, no. 5081, 1200-1205.

Matee, M. I. & R. Matre. 2001, "Pathogenic isolates in meningitis patients in Dar Es Salaam, Tanzania.", *East Afr. Med. J.*, vol. 78, no. 9, 458-460.

Mavroidi, A., D. Aanensen, D. Godoy, I. Skovsted, M. Kaltoft, P. Reeves, S. Bentley & B. Spratt. 2007, "Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci", *J. Bacteriol.*, vol. 189, no. 21, 7841-7855.

Mavroidi, A., D. Godoy, D. Aanensen, D. Robinson, S. Hollingshead & B. Spratt. 2004, "Evolutionary genetics of the capsular locus of serogroup 6 pneumococci", *J. Bacteriol.*, vol. 186, no. 24, 8181-8192.

McAllister, L. J., H.-J. Tseng, A. D. Ogunniyi, M. P. Jennings, A. G. McEwan & J. C. Paton. 2004, "Molecular analysis of the *psa* permease complex of *Streptococcus pneumoniae*.", *Mol. Microbiol.*, vol. 53, no. 3, 889-901.

McCool, T. L., T. R. Cate, G. Moy & J. N. Weiser. 2002, "The immune response to pneumococcal proteins during experimental human carriage", *J. Exp. Med.*, vol. 195, no. 3, 359-365.

McDougal, L. K., F. C. Tenover, L. N. Lee, J. K. Rasheed, J. E. Patterson, J. H. Jorgensen & D. J. LeBlanc. 1998, "Detection of Tn917-like sequences within a Tn916-like conjugative transposon (Tn3872) in erythromycin-resistant isolates of *Streptococcus pneumoniae*", *Antimicrob. Agents and Chemother.*, vol. 42, no. 9, 2312-2318.

McEllistrem, M. C., J. Adams, E. O. Mason & E. R. Wald. 2003, "Epidemiology of acute otitis media caused by *Streptococcus pneumoniae* before and after licensure of the 7-valent pneumococcal protein conjugate vaccine.", *J. Infect. Dis.*, vol. 188, no. 11, 1679-1684.

McEllistrem, M. C., A. C. Noller, S. Visweswaran, J. M. Adams & L. H. Harrison. 2004, "Serotype 14 variants of the France 9V(-3) clone from Baltimore, Maryland, can be differentiated by the *cpsB* gene", *J. Clin. Microbiol.,* vol. 42, no. 1, 250-256.

Mcgee, L., L. McDougal, J. Zhou, B. G. Spratt, F. C. Tenover, R. George, R. Hakenbeck, W. Hryniewicz, J. C. Lefévre, A. Tomasz & K. P. Klugman. 2001, "Nomenclature of major antimicrobial-resistant clones of *Streptococcus pneumoniae* defined by the pneumococcal molecular epidemiology network", *J. Clin. Microbiol.,* vol. 39, no. 7, 2565-2571.

McKee, C. M., M. B. Penno, M. Cowman, M. D. Burdick, R. M. Strieter, C. Bao & P. W. Noble. 1996, "Hyaluronan (HA) fragments induce chemokine gene expression in alveolar macrophages. The role of HA size and CD44", *J. Clin. Invest.,* vol. 98, no. 10, 2403-2413.

Meats, E., A. B. Brueggemann, M. C. Enright, K. Sleeman, D. T. Griffiths, D. W. Crook & B. G. Spratt. 2003, "Stability of serotypes during nasopharyngeal carriage of *Streptococcus pneumoniae*", *J. Clin. Microbiol.,* vol. 41, no. 1, 386-392.

Melegaro, A., Y. Choi, R. Pebody & N. Gay. 2007, "Pneumococcal carriage in United Kingdom families: estimating serotype-specific transmission parameters from longitudinal data", *Am. J. Epidemiol.,* vol. 166, no. 2, 228-235.

Melegaro, A. & W. J. Edmunds. 2004, "The 23-valent pneumococcal polysaccharide vaccine. Part I. Efficacy of PPV in the elderly: a comparison of meta-analyses.", *Eur. J. of Epidemiol.,* vol. 19, no. 4, 353-363.

Melin, M., H. Jarva, L. Siira, S. Meri, H. Kayhty & M. Väkeväinen. 2009, "*Streptococcus pneumoniae* capsular serotype 19F is more resistant to C3 deposition and less sensitive to opsonophagocytosis than serotype 6B.", *Infect. Immun.,* vol. 77, no. 2, 676-684.

Melin, M., K. Trzcinski, M. Antonio, S. Meri, R. Adegbola, T. Kaijalainen, H. Kayhty & M. Väkeväinen. 2010, "Serotype-related variation in susceptibility to complement deposition and opsonophagocytosis among clinical isolates of *Streptococcus pneumoniae*.", *Infect. Immun.,* vol. 78, no. 12, 5252-5261.

Mermin, J., J. Lule, J. P. Ekwaru, S. Malamba, R. Downing, R. Ransom, F. Kaharuza, D. Culver, F. Kizito, R. Bunnell, A. Kigozi, D. Nakanjako, W. Wafula & R. Quick. 2004, "Effect of co-trimoxazole prophylaxis on morbidity, mortality, CD4-cell count, and viral load in HIV infection in rural Uganda", *Lancet,* vol. 364, no. 9443, 1428-1434.

Messmer, T. O., C. M. Black & R. R. Facklam. 1995, "Discrimination of *Streptococcus pneumoniae* from other upper respiratory tract streptococci by arbitrarily primed PCR", *Clin. Biochem.*, vol. 28, no. 6, 567-572.

Mijakovic, I., S. Poncet, G. Boël, A. Mazé, S. Gillet, E. Jamet, P. Decottignies, C. Grangeassee, P. Doublet, P. Le Maréchal, J. Deutcher. 2003, "Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases", *EMBO J.*, vol. 22, no. 18, 4709-4718.

Millar, E. V., K. L. O'Brien, E. R. Zell, M. A. Bronsdon, R. Reid & M. Santosham. 2009, "Nasopharyngeal carriage of *Streptococcus pneumoniae* in Navajo and White Mountain Apache children before the introduction of pneumococcal conjugate vaccine.", *Pediatr. Infect. Dis. J.*, vol. 28, no. 8, 711-716.

Miller, E., N. J. Andrews, P. A. Waight, M. P. Slack & R. C. George. 2011, "Herd immunity and serotype replacement 4 years after seven-valent pneumococcal conjugate vaccination in England and Wales: an observational cohort study.", *Lancet Infect. Dis.*, vol. 11, no. 10, 760-768.

Mira, A., A. B. Martín-Cuadrado, G. D'Auria & F. Rodríguez-Valera. 2010, "The bacterial pan-genome:a new paradigm in microbiology", *Int. Microbiol.*, vol. 13, no. 2, 45-57.

Mitchell, A. M. & T. J. Mitchell. 2010, "*Streptococcus pneumoniae*: virulence factors and variation", *Clin. Microbiol. Infect.*, vol. 16, no. 5, 411-418.

Mitchell, T. J., P. W. Andrew, F. K. Saunders, A. N. Smith & G. J. Boulnois. 1991, "Complement activation and antibody binding by pneumolysin via a region of the toxin homologous to a human acute-phase protein", *Mol. Microbiol.*, vol. 5, no. 8, 1883-1888.

Mollerach, M., R. López & E. García. 1998, "Characterization of the *galU* gene of *Streptococcus pneumoniae* encoding a uridine diphosphoglucose pyrophosphorylase: a gene essential for capsular polysaccharide biosynthesis", *J. Exp. Med.*, vol. 188, no. 11, 2047-2056.

Montgomery, J. M., D. Lehmann, T. Smith, A. Michael, B. Joseph, T. Lupiwa, C. Coakley, V. Spooner, B. Best & I. D. Riley. 1990, "Bacterial colonization of the upper respiratory tract and its association with acute lower respiratory tract infections in Highland children of Papua New Guinea.", *Rev. Infect. Dis.*, vol. 12 Suppl 8, no. S1006-1016.

Moore, C. E., A. Sengduangphachanh, T. Thaojaikong, J. Sirisouk, D. Foster, R. Phetsouvanh, L. McGee, D. W. Crook, P. N. Newton & S. J. Peacock. 2010, "Enhanced determination of *Streptococcus pneumoniae* serotypes associated with invasive disease in Laos by using a real-time polymerase chain reaction serotyping assay with cerebrospinal fluid", *Am. J. Trop. Med. Hyg,* vol. 83, no. 3, 451-457.

Moore, M. R., R. E. Gertz, R. L. Woodbury, G. A. Barkocy-Gallagher, W. Schaffner, C. Lexau, K. Gershman, A. Reingold, M. Farley, L. H. Harrison, J. L. Hadler, N. M. Bennett, A. R. Thomas, L. McGee, T. Pilishvili, A. B. Brueggemann, C. G. Whitney, J. H. Jorgensen & B. Beall. 2008, "Population snapshot of emergent *Streptococcus pneumoniae* serotype 19A in the United States, 2005", *J. Infect. Dis.,* vol. 197, no. 7, 1016-1027.

Morais, L., M. d. G. Carvalho, A. Roca, B. Flannery, I. Mandomando, M. Soriano-Gabarró, B. Sigauque, P. Alonso & B. Beall. 2007, "Sequential multiplex PCR for identifying pneumococcal capsular serotypes from South-Saharan African clinical isolates", *J. Med. Microbiol.,* vol. 56, no. 9, 1181-1184.

Morand, B. & K. Mühlemann. 2007, "Heteroresistance to penicillin in *Streptococcus pneumoniae*", *Proc. Nat. Acad. Sci. U. S. A.,* vol. 104, no. 35, 14098-14103.

Moreno, A. T., M. L. S. Oliveira, D. M. Ferreira, P. L. Ho, M. Darrieux, L. C. C. Leite, J. M. C. Ferreira, F. C. Pimenta, A.-L. S. S. Andrade & E. N. Miyaji. 2010, "Immunization of mice with single PspA fragments induces antibodies capable of mediating complement deposition on different pneumococcal strains and cross-protection.", *Clin. Vaccine Immun,* vol. 17, no. 3, 439-446.

Moreno, J., E. Hernández, O. Sanabria & E. Castañeda. 2005, "Detection and serotyping of *Streptococcus pneumoniae* from nasopharyngeal samples by PCR-based multiplex assay", *J. Clin. Microbiol.,* vol. 43, no. 12, 6152-6154.

Morgan, P. J., S. C. Hyman, O. Byron, P. W. Andrew, T. J. Mitchell & A. J. Rowe. 1994, "Modeling the bacterial protein toxin, pneumolysin, in its monomeric and oligomeric form", *J. Biol. Chem.,* vol. 269, no. 41, 25315-25320.

Morona, J. K., D. C. Miller, R. Morona & J. C. Paton. 2004, "The effect that mutations in the conserved capsular polysaccharide biosynthesis genes *cpsA, cpsB,* and *cpsD* have on virulence of *Streptococcus pneumoniae*", *J. Infect.Dis.,* vol. 189, no. 10, 1905-1913.

Morona, J. K., R. Morona, D. C. Miller & J. C. Paton. 2002, "*Streptococcus pneumoniae* capsule biosynthesis protein CpsB is a novel manganese-

dependent phosphotyrosine-protein phosphatase", *J. Bacteriol.,* vol. 184, no. 2, 577-583.

Morona, J. K., R. Morona, D. C. Miller & J. C. Paton. 2003, "Mutational analysis of the carboxy-terminal (YGX)4 repeat domain of CpsD, an autophosphorylating tyrosine kinase required for capsule biosynthesis in *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 185, no. 10, 3009-3019.

Morona, J. K., R. Morona & J. C. Paton. 1997, "Characterization of the locus encoding the *Streptococcus pneumoniae* type 19F capsular polysaccharide biosynthetic pathway", *Mol. Microbiol.,* vol. 23, no. 4, 751-763.

Morona, J. K., R. Morona & J. C. Paton. 1999a, "Analysis of the 5' portion of the type 19A capsule locus identifies two classes of *cpsC, cpsD,* and *cpsE* genes in *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 181, no. 11, 3599-3605.

Morona, J. K., R. Morona & J. C. Paton. 1999b, "Comparative genetics of capsular polysaccharide biosynthesis in *Streptococcus pneumoniae* types belonging to serogroup 19", *J. Bacteriol.,* vol. 181, no. 17, 5355-5364.

Morona, J. K., R. Morona & J. C. Paton. 2006, "Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease", *Proc. Natl. Acad. Sci. U. S. A.,* vol. 103, no. 22, 8505-8510.

Morona, J. K., J. C. Paton, D. C. Miller & R. Morona. 2000, "Tyrosine phosphorylation of CpsD negatively regulates capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*", *Mol. Microbiol.,* vol. 35, no. 6, 1431-1442.

Moscoso, M. & E. García. 2009, "Transcriptional regulation of the capsular polysaccharide biosynthesis locus of streptococcus pneumoniae: a bioinformatic analysis", *DNA Res.,* vol. 16, no. 3, 177-186.

Moscoso, M., E. García & R. López. 2006, "Biofilm formation by *Streptococcus pneumoniae*: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion", *J. Bacteriol.,* vol. 188, no. 22, 7785-7795.

Moyo, S. J., M. Steinbakk, S. Aboud, N. Mkopi, M. Kasubi, B. Blomberg, K. Manji, E. F. Lyamuya, S. Y. Maselle & N. Langeland. 2012, "Penicillin resistance and serotype distribution of *Streptococcus pneumoniae* in nasopharyngeal carrier children under-5 years in Dar es Salaam, Tanzania.", *J. Med. Microbiol.,* doi: 10.1099/jmm.0.042598-0.

Mudhune, S., M. Wamae & N. S. f. P. D. i. t. E. A. Region. 2009, "Report on invasive disease and meningitis due to *Haemophilus influenzae* and *Streptococcus pneumonia* from the Network for Surveillance of Pneumococcal Disease in the East African Region.", *Clin. Infect. Dis.,* vol. 48 Suppl 2, no. S147-152.

Müller-Graf, C. D., A. M. Whatmore, S. J. King, K. Trzcinski, A. P. Pickerill, N. Doherty, J. Paul, D. Griffiths, D. Crook & C. G. Dowson. 1999, "Population biology of *Streptococcus pneumoniae* isolated from oropharyngeal carriage and invasive disease", *Microbiology,* vol. 145, no. 11, 3283-3293.

Muñoz, R., M. Mollerach & R. López. 1997, "Molecular organization of the genes required for the synthesis of type 1 capsular polysaccharide of *Streptococcus pneumoniae*: formation of binary encapsulated pneumococci and identification of cryptic dTDP-rhamnose biosynthesis genes.", *Mol. Microbiol.,* vol. 25, no. 1, 79-82.

Muñoz-Elías, E. J., J. Marcano & A. Camilli. 2008, "Isolation of *Streptococcus pneumoniae* biofilm mutants and their characterization during nasopharyngeal colonization", *Infect. Immun.,* vol. 76, no. 11, 5049-5061.

Murphy, T. F. 2006, "Otitis media, bacterial colonization, and the smoking parent", *Clin. Infect. Dis.,* vol. 42, no. 7, 904-906.

Muzzi, A. & C. Donati. 2011, "Population genetics and evolution of the pan-genome of *Streptococcus pneumoniae*", *Int.J. Med. Microbiol.,* vol. 301, no. 8, 619-622.

Nahm, M. H., J. Lin, J. A. Finkelstein & S. I. Pelton. 2009, "Increase in the prevalence of the newly discovered pneumococcal serotype 6C in the nasopharynx after introduction of pneumococcal conjugate vaccine", *J. Infect. Dis.* vol. 199, no. 3, 320-325.

Neeleman, C., C. H. W. Klaassen, D. M. Klomberg, H. A. de Valk & J. W. Mouton. 2004, "Pneumolysin is a key factor in misidentification of macrolide-resistant *Streptococcus pneumoniae* and is a putative virulence factor of *S. mitis* and other streptococci", *J. Clin. Microbiol.,* vol. 42, no. 9, 4355-4357.

Nelson, A., J. Ries, F. Bagnoli, S. Dahlberg, S. Fälker, S. Rounioja, J. Tschöp, E. Morfeldt, I. Ferlenghi & M. Hilleringmann. 2007a, "RrgA is a pilus- associated adhesin in *Streptococcus pneumoniae*", *Mol. Microbiol.,* vol. 66, no. 2, 329-340.

Nelson, A., A. Roche, J. Gould, K. Chim, A. Ratner & J. Weiser. 2007b, "Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance", *Infect. Immun.,* vol. 75, no. 1, 83-90.

Nesin, M., M. Ramirez & A. Tomasz. 1998, "Capsular transformation of a multidrug-resistant *Streptococcus pneumoniae* in vivo", *J. Infect. Dis.,* vol. 177, no. 3, 707-713.

Njanpop Lafourcade, B.-M., O. Sanou, M. van der Linden, N. Levina, M. Karanfil, S. Yaro, T. A. Tamekloe & J. E. Mueller. 2010, "Serotyping pneumococcal meningitis cases in the African meningitis belt by use of multiplex PCR with cerebrospinal fluid", *J. Clin. Microbiol.,* vol. 48, no. 2, 612-614.

Nolan, T., R. E. Hands, W. Ogunkolade & S. A. Bustin. 2006, "SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations", *Analytical biochemistry,* vol. 351, no. 2, 308-310.

Nuorti, J., J. Butler, J. Crutcher, R. Guevara, D. Welch, P. Holder & J. Elliott. 1998, "An outbreak of multidrug-resistant pneumococcal pneumonia and bacteremia among unvaccinated nursing home residents", *N. Engl. J. Med. England,* vol. 338, no. 26, 1861-1868.

O'Brien, K., M. Bronsdon, R. Dagan, P. Yagupsky, J. Janco, J. Elliott, C. Whitney, Y. Yang, L. Robinson & B. Schwartz. 2001, "Evaluation of a medium (STGG) for transport and optimal recovery of *Streptococcus pneumoniae* from nasopharyngeal secretions collected during field studies", *J. Clin. Microbiol.,* vol. 39, no. 3, 1021-1024.

O'Brien, K., E. V. Millar, E. R. Zell, M. Bronsdon, R. Weatherholtz, R. Reid, J. Becenti, S. Kvamme, C. G. Whitney & M. Santosham. 2007, "Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial", *J. Infect. Dis.,* vol. 196, no. 8, 1211-1220.

O'Brien, K., L. Moulton, R. Reid, R. Weatherholtz, J. Oski, L. Brown, G. Kumar, A. Parkinson, D. Hu & J. Hackell. 2003, "Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial", *Lancet,* vol. 362, no. 9381, 355-361.

O'Brien, K., H. Nohynek & W. H. O. P. V. T. C. W. Group. 2003, "Report from a WHO Working Group: standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*", *Pediatr. Infect. Dis. J.,* vol. 22, no. 2, e1-11.

O'Brien, K., L. Wolfson, J. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. Levine & T. Cherian. 2009, "Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates", *Lancet,* vol. 374, no. 9693, 893-902.

O'Brien, K. L., M. Hochman & D. Goldblatt. 2007, "Combined schedules of pneumococcal conjugate and polysaccharide vaccines: is hyporesponsiveness an issue?", *Lancet Infect. Dis.,* vol. 7, no. 9, 597-606.

O'Halloran D, M. & M. T. Cafferkey. 2005, "Multiplex PCR for identification of seven *Streptococcus pneumoniae* serotypes targeted by a 7-valent conjugate vaccine", *J. Clin. Microbiol.,* vol. 43, no. 7, 3487-3490.

Obaro, S. K., R. A. Adegbola, W. A. Banya & B. M. Greenwood. 1996, "Carriage of pneumococci after pneumococcal vaccination", *Lancet,* vol. 348, no. 9022, 271-272.

Obert, C., J. Sublett, D. Kaushal, E. Hinojosa, T. Barton, E. Tuomanen & C. Orihuela. 2006, "Identification of a candidate Streptococcus pneumoniae core genome and regions of diversity correlated with invasive pneumococcal disease", *Infect. Immun.,* vol. 74, no. 8, 4766-4777

Ogami, M., M. Hotomi, A. Togawa & N. Yamanaka. 2010, "A comparison of conventional and molecular microbiology in detecting differences in pneumococcal colonization in healthy children and children with upper respiratory illness.", *Eur. J. Pediatr.,* vol. 169, no. 10, 1221-1225.

Oggioni, M., C. Trappetti, A. Kadioglu, M. Cassone, F. Iannelli, S. Ricci, P. Andrew & G. Pozzi. 2006, "Switch from planktonic to sessile life: a major event in pneumococcal pathogenesis", *Mol. Microbiol.,* vol. 61, no. 5, 1196-1210.

Oggioni, M. R., F. Iannelli, S. Ricci, D. Chiavolini, R. Parigi, C. Trappetti, J.-P. Claverys & G. Pozzi. 2004, "Antibacterial activity of a competence-stimulating peptide in experimental sepsis caused by *Streptococcus pneumoniae*", *Antimicrob. Agents Chemother.,* vol. 48, no. 12, 4725-4732.

Oggioni, M. R. & G. Pozzi. 2001, "Comparative genomics for identification of clone-specific sequence blocks in *Streptococcus pneumoniae*", *FEMS Microbiol. Lett.,* vol. 200, no. 2, 137-143.

Orihuela, C. J., J. Mahdavi, J. Thornton, B. Mann, K. G. Wooldridge, N. Abouseada, N. J. Oldfield, T. Self, D. A. A. Ala'Aldeen & E. I. Tuomanen. 2009, "Laminin receptor initiates bacterial contact with the blood brain barrier in experimental meningitis models.", *J. Clin. Invest.,* vol. 119, no. 6, 1638-1646.

Oriyo, N. M. (2011). Epidemiology of *Streptococcus pneumoniae* in Northern Tanzania. Ph.D., University College London.

Örtqvist, Å. 2001, "Pneumococcal vaccination: current and future issues", *Eur. Respir. J.*, vol. 18, no. 1, 184-195.

Pai, R., R. E. Gertz & B. Beall. 2006, "Sequential multiplex PCR approach for determining capsular serotypes of Streptococcus pneumoniae isolates", *J. Clin. Microbiol.*, vol. 44, no. 1, 124-131.

Pai, R., J. Limor & B. Beall. 2005, "Use of pyrosequencing to differentiate *Streptococcus pneumoniae* serotypes 6A and 6B", *J. Clin. Microbiol.*, vol. 43, no. 9, 4820-4822.

Paolucci, M., M. P. Landini & V. Sambri. 2010, "Conventional and molecular techniques for the early diagnosis of bacteraemia", *Int. J. Antimicrob. Agents,* vol. 36 Suppl 2, no. S6-16.

Paradise, J. L., H. E. Rockette, D. K. Colborn, B. S. Bernard, C. G. Smith, M. Kurs-Lasky & J. E. Janosky. 1997, "Otitis media in 2253 Pittsburgh-area infants: prevalence and risk factors during the first two years of life.", *Pediatrics,* vol. 99, no. 3, 318-333.

Park, H. K., S.-J. Lee, J. W. Yoon, J. W. Shin, H.-S. Shin, J.-K. Kook, S. C. Myung & W. Kim. 2010, "Identification of the *cpsA* gene as a specific marker for the discrimination of *Streptococcus pneumoniae* from viridans group streptococci", *J. Med. Microbiol.,* vol. 59, no. 10, 1146-1152.

Park, I. H., S. Park, S. K. Hollingshead & M. H. Nahm. 2007, "Genetic basis for the new pneumococcal serotype, 6C", *Infect. Immun.,* vol. 75, no. 9, 4482-4489.

Park, M. K., D. E. Briles & M. H. Nahm. 2000, "A latex bead-based flow cytometric immunoassay capable of simultaneous typing of multiple pneumococcal serotypes (Multibead assay)", *Clin. Diagn. Lab. Immun.,* vol. 7, no. 3, 486-489.

Parker, D., G. Soong, P. Planet, J. Brower, A. J. Ratner & A. Prince. 2009, "The NanA neuraminidase of *Streptococcus pneumoniae* is involved in biofilm formation", *Infect. Immun.,* vol. 77, no. 9, 3722-3730.

Paton, J. C. & A. Ferrante. 1983, "Inhibition of human polymorphonuclear leukocyte respiratory burst, bactericidal activity, and migration by pneumolysin", *Infect. Immun.*, vol. 41, no. 3, 1212-1216.

Paton, J. C., B. Rowan-Kelly & A. Ferrante. 1984, "Activation of human complement by the pneumococcal toxin pneumolysin", *Infect. Immun.*, vol. 43, no. 3, 1085-1087.

Pearce, B. J., F. Iannelli & G. Pozzi. 2002, "Construction of new unencapsulated (rough) strains of *Streptococcus pneumoniae*", *Res. Microbiol.*, vol. 153, no. 4, 243-247.

Peerbooms, P. & M. Engelen. 2002, "Nasopharyngeal carriage of potential bacterial pathogens related to day care attendance, with special reference to the molecular epidemiology of *Haemophilus influenzae*", *J. Clin. Microbiol.*, vol. 40, no. 8, 2832-2836.

Pelton, S. I., A. M. Loughlin & C. D. Marchant. 2004, "Seven valent pneumococcal conjugate vaccine immunization in two Boston communities: changes in serotypes and antimicrobial susceptibility among *Streptococcus pneumoniae* isolates.", *Pediatr. Infect. Dis. J.*, vol. 23, no. 11, 1015-1022.

Perichon, B., J. Tankovic & P. Courvalin. 1997, "Characterization of a mutation in the *parE* gene that confers fluoroquinolone resistance in *Streptococcus pneumoniae", Antimicrob. Agents Chemother.*, vol. 41, no. 5, 1166-1167.

Pericone, C. D., D. Bae, M. Shchepetov, T. McCool & J. N. Weiser. 2002, "Short-sequence tandem and nontandem DNA repeats and endogenous hydrogen peroxide production contribute to genetic instability of *Streptococcus pneumoniae*.", *J. Bacteriol.*, vol. 184, no. 16, 4392-4399.

Pericone, C. D., K. Overweg, P. W. Hermans & J. N. Weiser. 2000, "Inhibitory and bactericidal effects of hydrogen peroxide production by *Streptococcus pneumoniae* on other inhabitants of the upper respiratory tract", *Infect. Immun.*, vol. 68, no. 7, 3990-3997.

Perry, J. A., D. G. Cvitkovitch & C. M. Lévesque. 2009, "Cell death in *Streptococcus mutans* biofilms: a link between CSP and extracellular DNA", *FEMS Microbiol. Lett.*, vol. 299, no. 2, 261-266.

Pestova, E. V., L. S. Håvarstein & D. A. Morrison. 1996, "Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system", *Mol. Microbiol.*, vol. 21, no. 4, 853-862.

Petersen, F. C., D. Pecharki & A. A. Scheie. 2004, "Biofilm mode of growth of *Streptococcus intermedius* favored by a competence-stimulating signaling peptide", *J. Bacteriol.,* vol. 186, no. 18, 6327-6331.

Peterson, S., R. T. Cline, H. Tettelin, V. Sharov & D. A. Morrison. 2000, "Gene expression analysis of the *Streptococcus pneumoniae* competence regulons by use of DNA microarrays", *J. Bacteriol.,* vol. 182, no. 21, 6192-6202.

Peterson, S. N., C. K. Sung, R. Cline, B. V. Desai, E. C. Snesrud, P. Luo, J. Walling, H. Li, M. Mintz, G. Tsegaye, P. C. Burr, Y. Do, S. Ahn, J. Gilbert, R. D. Fleischmann & D. A. Morrison. 2004, "Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays", *Mol. Microbiol.,* vol. 51, no. 4, 1051-1070.

Pettigrew, M. M., K. P. Fennie, M. P. York, J. Daniels & F. Ghaffar. 2006, "Variation in the presence of neuraminidase genes among *Streptococcus pneumoniae* isolates with identical sequence types.", *Infect. Immun.,* vol. 74, no. 6, 3360-3365.

Pettigrew, M. M., J. F. Gent, K. Revai, J. A. Patel & T. Chonmaitree. 2008, "Microbial interactions during upper respiratory tract infections.", *Emerg. Infect. Dis.,* vol. 14, no. 10, 1584-1591.

Pikis, A., J. M. Campos, W. J. Rodriguez & J. M. Keith. 2001, "Optochin resistance in *Streptococcus pneumoniae*: mechanism, significance, and clinical implications.", *J. Infect. Dis.,* vol. 184, no. 5, 582-590.

Pillai, D., D. Shahinas & A. Buzina. 2009, "Genome-wide dissection of globally emergent multi-drug resistant serotype 19A *Streptococcus pneumoniae*", *BMC Genomics,* vol. 10, 642.

Porat, N., U. Amit, N. Givon-Lavi, E. Leibovitz & R. Dagan. 2010, "Increasing importance of multidrug-resistant serotype 6A *Streptococcus pneumoniae* clones in acute otitis media in southern Israel", *Pediatr. Infect. Dis. J,* vol. 29, no. 2, 126-130.

Porat, N., D. Greenberg, N. Givon-Lavi, D. S. Shuval, R. Trefler, O. Segev, W. P. Hanage & R. Dagan. 2006, "The important role of nontypable *Streptococcus pneumoniae* international clones in acute conjunctivitis", *J. Infect. Dis.,* vol. 194, no. 5, 689-696.

Poulsen, K., J. Reinholdt & M. Kilian. 1996, "Characterization of the *Streptococcus pneumoniae* immunoglobulin A1 protease gene (*iga*) and its translation product", *Infect. Immun.,* vol. 64, no. 10, 3957-3966.

Pozzi, G., L. Masala, F. Iannelli, R. Manganelli, L. Havarstein, L. Piccoli, D. Simon & D. Morrison. 1996, "Competence for genetic transformation in encapsulated strains of *Streptococcus pneumoniae*: two allelic variants of the peptide pheromone", *J. Bacteriol.,* vol. 178, no. 20, 6087-6090.

Pracht, D., C. Elm, J. Gerber, S. Bergmann, M. Rohde, M. Seiler, K. S. Kim, H. F. Jenkinson, R. Nau & S. Hammerschmidt. 2005, "PavA of *Streptococcus pneumoniae* modulates adherence, invasion, and meningeal inflammation", *Infect. Immun.,* vol. 73, no. 5, 2680-2689.

Prudhomme, M., L. Attaiech, G. Sanchez, B. Martin & J.-P. Claverys. 2006, "Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*.", *Science,* vol. 313, no. 5783, 89-92.

Ramirez, M., D. A. Morrison & A. Tomasz. 1997, "Ubiquitous distribution of the competence related genes *comA* and *comC* among isolates of *Streptococcus pneumoniae*", *Microb. Drug Resist.,* vol. 3, no. 1, 39-52.

Rapola, S., E. Salo, P. Kiiski, M. Leinonen & A. K. Takala. 1997, "Comparison of four different sampling methods for detecting pharyngeal carriage of *Streptococcus pneumoniae* and *Haemophilus influenzae* in children", *J. Clin. Microbiol.,* vol. 35, no. 5, 1077-1079.

Raymond, J., I. Le Thomas, F. Moulin, A. Commeau, D. Gendrel & P. Berche. 2000, "Sequential colonization by *Streptococcus pneumoniae* of healthy children living in an orphanage.", *J. Infect. Dis.,* vol. 181, no. 6, 1983-1988.

Read, T. D., S. L. Salzberg, M. Pop, M. Shumway, L. Umayam, L. Jiang, E. Holtzapple, J. D. Busch, K. L. Smith, J. M. Schupp, D. Solomon, P. Keim & C. M. Fraser. 2002, "Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*", *Science,* vol. 296, no. 5575, 2028-2033.

Regev-Yochay, G., R. Dagan, M. Raz, Y. Carmeli, B. Shainberg, E. Derazne, G. Rahav & E. Rubinstein. 2004a, "Association between carriage of *Streptococcus pneumoniae* and *Staphylococcus aureus* in children", *JAMA.,* vol. 292, no. 6, 716-720.

Regev-Yochay, G., M. Raz, R. Dagan, N. Porat, B. Shainberg, E. Pinco, N. Keller & E. Rubinstein. 2004b, "Nasopharyngeal carriage of *Streptococcus pneumoniae* by adults and children in community and family settings", *Clin. Infct. Dis.*, vol. 38, no. 5, 632-639.

Reid, S., W. Hong, K. Dew, D. Winn, B. Pang, J. Watt, D. Glover, S. Hollingshead & W. Swords. 2009, "*Streptococcus pneumoniae* forms surface-attached communities in the middle ear of experimentally infected chinchillas", *J. Infect. Dis.*, vol. 199, no. 6, 786-794.

Reis, J., T. Palma, G. Ribeiro, R. Pinheiro, C. Ribeiro, S. Cordeiro, H. da Silva Filho, M. Moschioni, T. Thompson & B. Spratt. 2008, "Transmission of *Streptococcus pneumoniae* in an urban slum community", *J. Infect.*, vol. 57, no. 3, 204-213.

Ren, B., A. J. Szalai, O. Thomas, S. K. Hollingshead & D. E. Briles. 2003, "Both family 1 and family 2 PspA proteins can inhibit complement deposition and confer virulence to a capsular serotype 3 strain of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 71, no. 1, 75-85.

Rey, L. C., B. Wolf, J. L. Moreira, D. Milatovic, J. Verhoef & C. K. Farhat. 2002, "Antimicrobial susceptibility and serotypes of nasopharyngeal *Streptococcus pneumoniae* in children with pneumonia and in children with pneumonia and in children attending day-care centres in Fortaleza, Brazil", *Int. J. Antimicrob. Agents*, vol. 20, no. 2, 86-92.

Richter, S. S., K. P. Heilmann, S. E. Beekmann, N. J. Miller, C. L. Rice & G. V. Doern. 2005, "The molecular epidemiology of *Streptococcus pneumoniae* with quinolone resistance mutations", *Clin. Infect. Dis.*, vol. 40, no. 2, 225-235.

Rijkers, G. T., E. A. Sanders, M. A. Breukels & B. J. Zegers. 1998, "Infant B cell responses to polysaccharide determinants.", *Vaccine*, vol. 16, no. 14-15, 1396-1400.

Rimini, R., B. Jansson, G. Feger, T. C. Roberts, M. de Francesco, A. Gozzi, F. Faggioni, E. Domenici, D. M. Wallace, N. Frandsen & A. Polissi. 2000, "Global analysis of transcription kinetics during competence development in *Streptococcus pneumoniae* using high density DNA arrays", *Mol. Microbiol.*, vol. 36, no. 6, 1279-1292.

Ring, A., J. N. Weiser & E. I. Tuomanen. 1998, "Pneumococcal trafficking across the blood-brain barrier. Molecular analysis of a novel bidirectional pathway", *J Clin. Invest.*, vol. 102, no. 2, 347-360.

Rivera-Olivero, I., M. Blommaart, D. Bogaert, P. Hermans & J. de Waard. 2009, "Multiplex PCR reveals a high rate of nasopharyngeal pneumococcal 7-valent conjugate vaccine serotypes co-colonizing indigenous Warao children in Venezuela", *J. Med. Microbiol.*, vol. 58, no. 5, 584-587.

Robinson, D. A., D. E. Briles, M. J. Crain & S. K. Hollingshead. 2002, "Evolution and virulence of serogroup 6 pneumococci on a global scale", *J. Bacteriol.*, vol. 184, no. 22, 6367-6375.

Robinson, D. A., K. M. Edwards, K. B. Waites, D. E. Briles, M. J. Crain & S. K. Hollingshead. 2001, "Clones of *Streptococcus pneumoniae* isolated from nasopharyngeal carriage and invasive disease in young children in central Tennessee", *J. Infect. Dis.*, vol. 183, no. 10, 1501-1507.

Robson, R. L., S. Essengue, N. A. Reed & R. T. Horvat. 2007, "Optochin resistance in *Streptococcus pneumoniae* induced by frozen storage in glycerol", *Diagn. Microbiol. and Infect. Dis.*, vol. 58, no. 2, 185-190.

Rodriguez-Barradas, M. C., R. A. Tharapel, J. E. Groover, K. P. Giron, C. E. Lacke, E. D. Houston, R. J. Hamill, M. C. Steinhoff & D. M. Musher. 1997, "Colonization by *Streptococcus pneumoniae* among human immunodeficiency virus-infected adults: prevalence of antibiotic resistance, impact of immunization, and characterization by polymerase chain reaction with BOX primers of isolates from persistent *S. pneumoniae* carriers", *J. Infect. Dis.*, vol. 175, no. 3, 590-597.

Romero-Steiner, S., D. Libutti, L. B. Pais, J. Dykes, P. Anderson, J. C. Whitin, H. L. Keyserling & G. M. Carlone. 1997, "Standardization of an opsonophagocytic assay for the measurement of functional antibody activity against *Streptococcus pneumoniae* using differentiated HL-60 cells", *Clin. Diagn. Lab. Immun.*, vol. 4, no. 4, 415-422.

Romero-Steiner, S., T. Pilishvili, J. S. Sampson, S. E. Johnson, A. Stinson, G. M. Carlone & E. W. Ades. 2003, "Inhibition of pneumococcal adherence to human nasopharyngeal epithelial cells by anti-PsaA antibodies", *Clin. Diagn. Lab. Immun.*, vol. 10, no. 2, 246-251.

Rosén, C., P. Christensen, B. Hovelius & K. Prellner. 1984, "A longitudinal study of the nasopharyngeal carriage of pneumococci as related to pneumococcal vaccination in children attending day-care centres.", *Acta. Otolaryngol.*, vol. 98, no. 5-6, 524-532.

Rosenow, C., P. Ryan, J. N. Weiser, S. Johnson, P. Fontan, A. Ortqvist & H. R. Masure. 1997, "Contribution of novel choline-binding proteins to adherence,

colonization and immunogenicity of *Streptococcus pneumoniae*", *Mol. Microbiol.*, vol. 25, no. 5, 819-829.

Rossjohn, J., R. J. Gilbert, D. Crane, P. J. Morgan, T. J. Mitchell, A. J. Rowe, P. W. Andrew, J. C. Paton, R. K. Tweten & M. W. Parker. 1998, "The molecular mechanism of pneumolysin, a virulence factor from *Streptococcus pneumoniae*", *J. Mol. Biol.*, vol. 284, no. 2, 449-461.

Ruoff, K. L. & A. L. Bisno (2010). Classification of Streptococci. Principles and Practices of Infectious Diseases. G. L. Mandell and J. E. Bennett. Philadelphia, Churchill Livingstone Elsevier**:** 2591-2592.

Royal Free Hampstead NHS Trust. (2011a). Investigation of Blood Culture Samples. Standard Operating Procudure. MIC.LP-GBacInvest-08.

Royal Free Hampstead NHS Trust. (2011b). Investigation of Fluid Samples. Standard Operating Procudure. MIC.LP-GBacInvest-15.

Royal Free Hampstead NHS Trust (2011c). Investigation of Respiratory Samples. Standard Operating Procudure. MIC.LP-GBacInvest-12.

Rubin, L. & A. Rizvi. 2004, "PCR-based assays for detection of *Streptococcus pneumoniae* serotypes 3, 14, 19F and 23F in respiratory specimens", *J. Med. Microbiol.*, vol. 53, no. 7, 595.

Rubins, J. B., T. J. Mitchell, P. W. Andrew & D. E. Niewoehner. 1994, "Pneumolysin activates phospholipase A in pulmonary artery endothelial cells", *Infect. Immun.*, vol. 62, no. 9, 3829-3836.

Rubins, J. B., A. H. Paddock, D. Charboneau, A. M. Berry, J. C. Paton & E. N. Janoff. 1998, "Pneumolysin in pneumococcal adherence and colonization", *Microb. Pathog.*, vol. 25, no. 6, 337-342.

Rukke, H. V., I. K. Hegna & F. C. Petersen. 2012, "Identification of a functional capsule locus in *Streptococcus mitis*.", *Mol. Oral Microbiol.*, vol. 27, no. 2, 95-108.

Russell, H., R. R. Facklam, J. F. Padula & R. Cooksey. 1978, "Capillary precipitin typing of *Streptococcus pneumoniae*.", *J. Clin. Microbiol.*, vol. 8, no. 4, 355-359.

Russell, H., J. A. Tharpe, D. E. Wells, E. H. White & J. E. Johnson. 1990, "Monoclonal antibody recognizing a species-specific protein from *Streptococcus pneumoniae*", *J. Clin. Microbiol.*, vol. 28, no. 10, 2191-2195.

Sá-Leão, R., S. Nunes, A. Brito-Avô, C. R. Alves, J. A. Carriço, J. Saldanha, J. S. Almeida, I. Santos-Sanches & H. de Lencastre. 2008, "High rates of transmission of and colonization by *Streptococcus pneumoniae* and Haemophilus influenzae within a day care center revealed in a longitudinal study", *J. Clin. Microbiol.*, vol. 46, no. 1, 225-234.

Sa-Leao, R., S. Nunes, A. Brito-Avô, N. Frazao, A. S. Simões, M. I. Crisóstomo, A. C. S. Paulo, J. Saldanha, I. Santos-Sanches & H. De Lencastre. 2009, "Changes in pneumococcal serotypes and antibiotypes carried by vaccinated and unvaccinated day-care centre attendees in Portugal, a country with widespread use of the seven-valent pneumococcal conjugate vaccine.", *Clin. Microbiol. Infect.*, vol. 15, no. 11, 1002-1007.

Sá-Leão, R., A. S. Simões, S. Nunes, N. G. Sousa, N. Frazão & H. de Lencastre. 2006, "Identification, prevalence and population structure of non-typable *Streptococcus pneumoniae* in carriage samples isolated from preschoolers attending day-care centres", *Microbiology*, vol. 152, no. 2, 367-376.

Sa-Leao, R., A. Tomasz, I. Santos Sanches & H. De Lencastre. 2002, "Pilot study of the genetic diversity of the pneumococcal nasopharyngeal flora among children attending day care centers", *J. Clin. Microbiol.*, vol. 40, no. 10, 3577-3585.

Saha, S., G. Darmstadt, A. Baqui, B. Hossain, M. Islam, D. Foster, H. Al-Emran, A. Naheed, S. El Arifeen & S. Luby. 2008, "Identification of serotype in culture negative pneumococcal meningitis using sequential multiplex PCR: implication for surveillance and vaccine design", *PLoS ONE*, vol. 3, no. 10, e3576.

Sampson, J. S., S. P. O'Connor, A. R. Stinson, J. A. Tharpe & H. Russell. 1994, "Cloning and nucleotide sequence analysis of *psaA*, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein homologous to previously reported *Streptococcus* sp. adhesins", *Infect. Immun.*, vol. 62, no. 1, 319-324.

Sanchez, C. J., B. J. Hurtgen, A. Lizcano, P. Shivshankar, G. T. Cole & C. J. Orihuela. 2011, "Biofilm and planktonic pneumococci demonstrate disparate immunoreactivity to human convalescent sera", *BMC Microbiology*, vol. 11, no. 1, 245.

Sandgren, A., K. Sjostrom, B. Olsson-Liljequist, B. Christensson, A. Samuelsson, G. Kronvall & B. Henriques Normark. 2004, "Effect of clonal and serotype-specific properties on the invasive capacity of *Streptococcus pneumoniae*", *J. Infect. Dis.,* vol. 189, no. 5, 785-796.

Schaub, I. G. & R. D. Reid. 1938, "A microscopic precipitin test for typing pneumococci ", *JAMA,* vol. 111, no. 14, 1285-1286.

Scott, D. A., S. F. Komjathy, B. T. Hu, S. Baker, L. A. Supan, C. A. Monahan, W. Gruber, G. R. Siber & S. P. Lockhart. 2007, "Phase 1 trial of a 13-valent pneumococcal conjugate vaccine in healthy adults", *Vaccine,* vol. 25, no. 33, 6164-6166.

Scott, J. R., W. P. Hanage, M. Lipsitch, E. V. Millar, L. H. Moulton, J. Hinds, M. Santosham & K. L. O'Brien. 2012, "Pneumococcal sequence type replacement among American Indian children : a comparison of pre- and routine-PCV7 eras", *Vaccine,* vol. 30, no. 13, 2376-2381.

Severin, A. & A. Tomasz. 1996, "Naturally occurring peptidoglycan variants of *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 178, no. 1, 168-174.

Shahinas, D., G. S. Tamber, G. Arya, A. Wong, R. Lau, F. Jamieson, J. H. Ma, D. C. Alexander, D. E. Low & D. R. Pillai. 2011, "Whole-Genome Sequence of *Streptococcus pseudopneumoniae* Isolate IS7493", *J. Bacteriol.,* vol. 193, no. 21, 6102-6103.

Shakhnovich, E., S. King & J. Weiser. 2002, "Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract", *Infect. Immun.,* vol. 70, no. 12, 7161-7164.

Shaper, M., S. K. Hollingshead, W. H. Benjamin & D. E. Briles. 2004, "PspA protects *Streptococcus pneumoniae* from killing by apolactoferrin, and antibody to PspA enhances killing of pneumococci by apolactoferrin [corrected]", *Infect. Immun.,* vol. 72, no. 9, 5031-5040.

Shayegani, M., L. M. Parsons, W. E. Gibbons & D. Campbell. 1982, "Characterization of nontypable *Streptococcus pneumoniae*-like organisms isolated from outbreaks of conjunctivitis", *J. Clin. Microbiol.,* vol. 16, no. 1, 8-14.

Shen, K., P. Antalis, J. Gladitz, S. Sayeed, A. Ahmed, S. Yu, J. Hayes, S. Johnson, B. Dice, R. Dopico, R. Keefe, B. Janto, W. Chong, J. Goodwin, R. M. Wadowsky, G. Erdos, J. C. Post, G. D. Ehrlich & F. Z. Hu. 2005, "Identification, distribution, and expression of novel genes in 10 clinical isolates of nontypeable *Haemophilus influenzae*", *Infect. Immun.*, vol. 73, no. 6, 3479-3491.

Shen, K., J. Gladitz, P. Antalis, B. Dice, B. Janto, R. Keefe, J. Hayes, A. Ahmed, R. Dopico, N. Ehrlich, J. Jocz, L. Kropp, S. Yu, L. Nistico, D. P. Greenberg, K. Barbadora, R. A. Preston, J. C. Post, G. D. Ehrlich & F. Z. Hu. 2006, "Characterization, distribution, and expression of novel genes among eight clinical isolates of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 74, no. 1, 321-330.

Shouval, D. S., N. Porat, R. Dagan, N. Keller, E. Bilavsky, Y. Sivan & J. Amir. 2010, "Bacteremia caused by a highly-resistant *Streptococcus pneumoniae* serotype 19A circulating in a daycare center", *Int. J. Infect. Dis.*, vol. 14 Suppl 3, no. e253-255.

Siberry, G., K. Brahmadathan, R. Pandian, M. Lalitha, M. Steinhoff & T. John. 2001, "Comparison of different culture media and storage temperatures for the long-term preservation of *Streptococcus pneumoniae* in the tropics", *Bull. World Health Organ.*, vol. 79, no. 1, 43-47.

Sibold, C., J. Henrichsen, A. König, C. Martin, L. Chalkley & R. Hakenbeck. 1994, "Mosaic pbpX genes of major clones of penicillin-resistant *Streptococcus pneumoniae* have evolved from *pbpX* genes of a penicillin-sensitive *Streptococcus oralis*", *Mol. Microbiol.*, vol. 12, no. 6, 1013-1023.

Silva, N. A., J. McCluskey, J. M. C. Jefferies, J. Hinds, A. Smith, S. C. Clarke, T. J. Mitchell & G. K. Paterson. 2006, "Genomic diversity between strains of the same serotype and multilocus sequence type among pneumococcal clinical isolates", *Infect. Immun.*, vol. 74, no. 6, 3513-3518.

Simell, B., M. Korkeila, H. Pursiainen, T. M. Kilpi & H. Käyhty. 2001, "Pneumococcal carriage and otitis media induce salivary antibodies to pneumococcal surface adhesin a, pneumolysin, and pneumococcal surface protein a in children", *J. Infect. Dis.*, vol. 183, no. 6, 887-896.

Simões, A. S., L. Pereira, S. Nunes, A. Brito-Avô, H. de Lencastre & R. Sá-Leão. 2011a, "Clonal evolution leading to maintenance of antibiotic resistance rates among colonizing pneumococci in the PCV7 era in Portugal.", *J. Clin. Microbiol.*, vol. 49, no. 8, 2810-2817.

Simões, A. S., R. Sá-Leão, M. J. Eleveld, D. A. Tavares, J. A. Carriço, H. J. Bootsma & P. W. M. Hermans. 2010, "Highly penicillin-resistant multidrug-resistant pneumococcus-like strains colonizing children in Oeiras, Portugal: genomic characteristics and implications for surveillance", *J. Clin. Microbiol.*, vol. 48, no. 1, 238-246.

Simões, A. S., C. Valente, H. de Lencastre & R. Sá-Leão. 2011b, "Rapid identification of noncapsulated *Streptococcus pneumoniae* in nasopharyngeal samples allowing detection of co-colonization and reevaluation of prevalence", *Diagn. Microbiol. Infect. Dis.*, vol. 71, no. 3, 208-216.

Singleton, R. J., T. W. Hennessy, L. R. Bulkow, L. L. Hammitt, T. Zulz, D. A. Hurlburt, J. C. Butler, K. Rudolph & A. Parkinson. 2007, "Invasive pneumococcal disease caused by nonvaccine serotypes among Alaska native children with high levels of 7-valent pneumococcal conjugate vaccine coverage", *JAMA*, vol. 297, no. 16, 1784-1792.

Sistek, V., M. Boissinot, D. K. Boudreau, A. Huletsky, F. J. Picard & M. G. Bergeron. 2011, "Development of a real-time PCR assay for the specific detection and identification of *Streptococcus pseudopneumoniae* using the *recA* gene", *Clin. Microbiol. Infect.*, doi: 10.1111/j.1469-0691.2011.03684.

Slotved, H.-C., M. Kaltoft, I. C. Skovsted, M. B. Kerrn & F. Espersen. 2004, "Simple, rapid latex agglutination test for serotyping of pneumococci (Pneumotest-Latex)", *J. Clin. Microbiol.*, vol. 42, no. 6, 2518-2522.

Sluijter, M., H. Faden, R. de Groot, N. Lemmens, W. H. Goessens, A. Van Belkum & P. W. Hermans. 1998, "Molecular characterization of pneumococcal nasopharynx isolates collected from children during their first 2 years of life", *J. Clin. Microbiol.*, vol. 36, no. 8, 2248-2253.

Smart, L. E. 1986, "Serotyping of *Streptococcus pneumoniae* strains by coagglutination.", *J. Clin. Pathol.*, vol. 39, no. 3, 328-331.

Smart, L. E., A. J. Dougall & R. W. Girdwood. 1987, "Identification of non-capsulate strains of *Streptococcus pneumoniae* by coagglutination.", *J. Clin. Pathol.*, vol. 40, no. 2, 243.

Smith, A. M. & K. P. Klugman. 1998, "Alterations in PBP 1A essential-for high-level penicillin resistance in *Streptococcus pneumoniae*", *Antimicrob. Agents Chemother.*, vol. 42, no. 6, 1329-1333.

Smith, H. E., M. Damman, J. van der Velde, F. Wagenaar, H. J. Wisselink, N. Stockhofe-Zurwieden & M. A. Smits. 1999, "Identification and characterization of the *cps* locus of *Streptococcus suis* serotype 2: the capsule protects against phagocytosis and is an important virulence factor.", *Infect. Immun.*, vol. 67, no. 4, 1750-1756.

Smith, S. K. & J. A. Washington. 1984, "Evaluation of the Pneumoslide latex agglutination test for identification of *Streptococcus pneumoniae*", *J. Clin. Microbiol.*, vol. 20, no. 3, 592-593.

Soewignjo, S., B. D. Gessner, A. Sutanto, M. Steinhoff, M. Prijanto, C. Nelson, A. Widjaya & S. Arjoso. 2001, "*Streptococcus pneumoniae* nasopharyngeal carriage prevalence, serotype distribution, and resistance patterns among children on Lombok Island, Indonesia.", *Clin. Infect. Dis.*, vol. 32, no. 7, 1039-1043.

Soininen, A., H. Pursiainen, T. Kilpi & H. Käyhty. 2001, "Natural development of antibodies to pneumococcal capsular polysaccharides depends on the serotype: association with pneumococcal carriage and acute otitis media in young children", *J. Infect. Dis.*, vol. 184, no. 5, 569-576.

Sondag, J. E., Morgens, R. K., Hoppe, J. E. & Marr, J. J. 1977, "Detection of pneumococci in respiratory specimens: clinical evaluation of gentamicin blood agar", *J. Clin. Microbiol.*, vol. 5, no. 4, 397-400.

Song, J., S. Jung, K. Ko & N. Kim. 2004, "High prevalence of antimicrobial resistance among clinical *Streptococcus pneumoniae* isolates in Asia (an ANSORP study)", *Antimicrob. Agents Chemother.*, vol. 48, no. 6, 2101-2107.

Song, J.-H., J. Y. Baek & K. S. Ko. 2011, "Comparison of capsular genes of *Streptococcus pneumoniae* serotype 6A, 6B, 6C, and 6D Isolates", *J. Clin. Microbiol.*, vol. 49, no. 5, 1758-1764.

Sørensen, U. B. 1993, "Typing of pneumococci by using 12 pooled antisera", *J. Clin. Microbiol.*, vol. 31, no. 8, 2097-2100.

Sørensen, U. B. & J. Henrichsen. 1987, "Cross-reactions between pneumococci and other streptococci due to C polysaccharide and F antigen", *J. Clin. Microbiol.*, vol. 25, no. 10, 1854-1859.

Sørensen, U. B., J. Henrichsen, H. C. Chen & S. C. Szu. 1990, "Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of

Streptococcus pneumoniae revealed by immunochemical methods", *Microb. Pathog.,* vol. 8, no. 5, 325-334.

Sridharan, G., T. J. John, M. K. Lalitha, L. H. Harrison & M. C. Steinhoff. 1994, "Serotypes of *Streptococcus pneumoniae* causing meningitis in southern India. Use of new direct latex agglutination antigen detection tests in cerebrospinal fluid.", *Diagn. Microbiol. Infect. Dis.,* vol. 18, no. 4, 211-214.

St. Sauver, J., C. F. Marrs, B. Foxman, P. Somsel, R. Madera & J. R. Gilsdorf. 2000, "Risk factors for otitis media and carriage of multiple strains of *Haemophilus influenzae* and *Streptococcus pneumoniae*", *Emerg. Infect. Dis.,* vol. 6, no. 6, 622-630.

Stanhope, M. J., S. L. Walsh, J. A. Becker, M. J. Italia, K. A. Ingraham, M. N. Gwynn, T. Mathie, J. A. Poupard, L. A. Miller, J. R. Brown & H. Amrine-Madsen. 2005, "Molecular evolution perspectives on intraspecific lateral DNA transfer of topoisomerase and gyrase loci in *Streptococcus pneumoniae*, with implications for fluoroquinolone resistance development and spread", *Antimicrob. Agents Chemother.,* vol. 49, no. 10, 4315-4326.

Stanhope, M. J., S. L. Walsh, J. A. Becker, L. A. Miller, T. Lefébure, P. Lang, P. D. P. Bitar & H. Amrine-Madsen. 2007, "The relative frequency of intraspecific lateral gene transfer of penicillin binding proteins 1a, 2b, and 2x, in amoxicillin resistant *Streptococcus pneumoniae*", *Infect. Genet. Evol.,* vol. 7, no. 4, 520-534.

Steinmoen, H., E. Knutsen & L. S. Håvarstein. 2002, "Induction of natural competence in *Streptococcus pneumoniae* triggers lysis and DNA release from a subfraction of the cell population", *Proc. Nat. Acad. Sci. U. S. A.,* vol. 99, no. 11, 7681-7686.

Steinmoen, H., A. Teigen & L. S. Håvarstein. 2003, "Competence-induced cells of *Streptococcus pneumoniae* lyse competence-deficient cells of the same strain during cocultivation", *J. Bacteriol.,* vol. 185, no. 24, 7176-7183.

Sung, R. Y., J. M. Ling, S. M. Fung, S. J. Oppenheimer, D. W. Crook, J. T. Lau & A. F. Cheng. 1995, "Carriage of *Haemophilus influenzae* and *Streptococcus pneumoniae* in healthy Chinese and Vietnamese children in Hong Kong", *Acta. Paediatr.,* vol. 84, no. 11, 1262-1267.

Suzuki, N., M. Seki, Y. Nakano, Y. Kiyoura, M. Maeno & Y. Yamashita. 2005, "Discrimination of *Streptococcus pneumoniae* from viridans group streptococci by genomic subtractive hybridization", *J. Clin. Microbiol.,* vol. 43, no. 9, 4528-4534.

Swaminathan, A., A. Mandlik, A. Swierczynski, A. Gaspar, A. Das & H. Ton-That. 2007, "Housekeeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*", *Mol Microbiol,* vol. 66, no. 4, 961-974.

Swiatlo, E., F. Champlin, S. Holman, W. Wilson & J. Watt. 2002, "Contribution of choline-binding proteins to cell surface properties of *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 70, no. 1, 412-415.

Syrjänen, R. K., T. M. Kilpi, T. H. Kaijalainen, E. E. Herva & A. K. Takala. 2001, "Nasopharyngeal carriage of *Streptococcus pneumoniae* in Finnish children younger than 2 years old", *J. Infect. Dis.*, vol. 184, no. 4, 451-459.

Talkington, D. F., D. L. Crimmins, D. C. Voellinger, J. Yother & D. E. Briles. 1991, "A 43-kilodalton pneumococcal surface protein, PspA: isolation, protective abilities, and structural analysis of the amino-terminal sequence.", *Infect. Immun.,* vol. 59, no. 4, 1285-1289.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei & S. Kumar. 2011, "MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods", *Mol. Biol. Evol.,* vol. 28, no. 10, 2731-2739.

Tankovic, J., B. Perichon, J. Duval & P. Courvalin. 1996, "Contribution of mutations in *gyrA* and *parC* genes to fluoroquinolone resistance of mutants of *Streptococcus pneumoniae* obtained *in vivo* and *in vitro*", *Antimicrob. Agents Chemother.,* vol. 40, no. 11, 2505-2510.

Tanzania Volunteers (2011). "Huruma Hospital." 2011, from http://www.tanzaniavolunteers.com/placements/huruma-hospital.

Tapiainen, T., T. Kujala, T. Kaijalainen, I. Ikäheimo, A. Saukkoriipi, M. Renko, J. Salo, M. Leinonen & M. Uhari. 2010, "Biofilm formation by *Streptococcus pneumoniae* isolates from paediatric patients", *APMIS,* vol. 118, no. 4, 255-260.

Tarragó, D., A. Fenoll, D. Sánchez-Tatay, L. A. Arroyo, C. Muñoz-Almagro, C. Esteva, W. P. Hausdorff, J. Casal & I. Obando. 2008, "Identification of pneumococcal serotypes from culture-negative clinical specimens by novel real-time PCR", *Clin. Microbiol. Infect.,* vol. 14, no. 9, 828-834.

Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy, T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C. R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J. Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J. Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G. Dimitrov, K. Watkins, K. J. B. O'Connor, S. Smith, T. R. Utterback, O. White, C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R. Wessels, R. Rappuoli & C. M. Fraser. 2005, "Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial 'pan-genome';" *Proc. Nat. Acad. Sci. U. S. A,* vol. 102, no. 39, 13950-13955.

Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J. Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F. Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M. R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L. Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey, I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A. Morrison, S. K. Hollingshead & C. M. Fraser. 2001, "Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*", *Science,* vol. 293, no. 5529, 498-506.

Thriemer, K., B. Ley, S. Ame, L. von Seidlein, G. D. Pak, N. Y. Chang, R. Hashim, W. H. Schmied, C. J.-L. Busch, S. Nixon, A. Morrissey, M. K. Puri, M. Ali, R. L. Ochiai, T. Wierzba, M. S. Jiddawi, J. D. Clemens, S. M. Ali & J. L. Deen. 2012, "The burden of invasive bacterial infections in Pemba, Zanzibar.", *PLoS ONE*, vol. 7, no. 2, e30350.

Tocheva, A. S., J. M. C. Jefferies, M. Christodoulides, S. N. Faust & S. C. Clarke. 2010, "Increase in serotype 6C pneumococcal carriage, United Kingdom", *Emerg. Infect. Dis.,* vol. 16, no. 1, 154-155.

Tocheva, A. S., J. M. C. Jefferies, H. Rubery, J. Bennett, G. Afimeke, J. Garland, M. Christodoulides, S. N. Faust & S. C. Clarke. 2011, "Declining serotype coverage of new pneumococcal conjugate vaccines relating to the carriage of *Streptococcus pneumoniae* in young children", *Vaccine,* vol. 29, no. 26, 4400-4404.

Tokuriki, N. & D. S. Tawfik. 2009, "Protein dynamism and evolvability", *Science,* vol. 324, no. 5924, 203-207.

Tomasson, G., T. Gudnason & K. G. Kristinsson. 2005, "Dynamics of pneumococcal carriage among healthy Icelandic children attending day-care centres.", *Scand. J. Infect. Dis.,* vol. 37, no. 6-7, 422-428.

Tomasz, A. 1967, "Choline in the cell wall of a bacterium: novel type of polymer-linked choline in pneumococcus", *Science,* vol. 157, no. 789, 694-697.

Tomasz, A. 1968, "Biological consequences of the replacement of choline by ethanolamine in the cell wall of pneumococcus: chanin formation, loss of transformability, and loss of autolysis", *Proc. Natl. Acad. Sci. U. S. A.,* vol. 59, no. 1, 86-93.

Tomasz, A. & R. D. Hotchkiss. 1964, "Regulation of the transformability of pneumococcal cultures by macromolecular cell products", *Proc. Nat. Acad. Sci U. S. A.,* vol. 51, no. 480-487.

Tomasz, A., P. Moreillon & G. Pozzi. 1988, "Insertional inactivation of the major autolysin gene of *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 170, no. 12, 5931-5934.

Tomasz, A. & M. Westphal. 1971, "Abnormal autolytic enzyme in a pneumococus with altered teichoic acid composition", *Proc. Nat. Acad. Sci. U. S. A.,* vol. 68, no. 11, 2627-2630.

Tomita, Y., A. Okamoto, K. Yamada, T. Yagi, Y. Hasegawa & M. Ohta. 2011a, "A new microarray system to detect *Streptococcus pneumoniae* serotypes", *J. Biomed. Biotech.,* vol. 2011, doi: 10.1155/2011/352736.

Tomita, Y., Y. Tomita, Okamoto, Yamada, Yagi, Hasegawa & Ohta. 2011b, "Capsular-type prediction by phylogenetic tree of glycosyltransferase gene polymorphism in *Streptococcus pneumoniae*", *OAB.,* vol. 3, 67-73.

Tong, H. H., M. A. McIver, L. M. Fisher & T. F. DeMaria. 1999, "Effect of lacto-N-neotetraose, asialoganglioside-GM1 and neuraminidase on adherence of otitis media-associated serotypes of *Streptococcus pneumoniae* to chinchilla tracheal epithelium", *Microb. Pathog.,* vol. 26, no. 2, 111-119.

Tortosa, P. & D. Dubnau. 1999, "Competence for transformation: a matter of taste", *Curr. Opin. Microbiol.,* vol. 2, no. 6, 588-592.

Traore, Y., T. A. Tameklo, B.-M. Njanpop Lafourcade, M. Lourd, S. Yaro, D. Niamba, A. Drabo, J. E. Mueller, J.-L. Koeck & B. D. Gessner. 2009, "Incidence, seasonality, age distribution, and mortality of pneumococcal meningitis in Burkina Faso and Togo.", *Clin. Infect. Dis.,* vol. 48 Suppl 2, no. S181-189.

Trappetti, C., L. Gualdi, L. Di Meola, P. Jain, C. C. Korir, P. Edmonds, F. Iannelli, S. Ricci, G. Pozzi & M. R. Oggioni. 2011a, "The impact of the competence quorum sensing system on *Streptococcus pneumoniae* biofilms varies depending on the experimental model", *BMC Microbiology,* vol. 11, no. 1, 75.

Trappetti, C., A. Kadioglu, M. Carter, J. Hayre, F. Iannelli, G. Pozzi, P. W. Andrew & M. R. Oggioni. 2009, "Sialic acid: a preventable signal for pneumococcal biofilm formation, colonization, and invasion of the host", *J. Infect. Dis.,* vol. 199, no. 10, 1497-1505.

Trappetti, C., A. D. Ogunniyi, M. R. Oggioni & J. C. Paton. 2011b, "Extracellular matrix formation enhances the ability of *Streptococcus pneumoniae* to cause invasive disease.", *PLoS ONE,* vol. 6, no. 5, e19844.

Trappetti, C., A. J. Potter, A. W. Paton, M. R. Oggioni & J. C. Paton. 2011c, "LuxS mediates iron-dependent biofilm formation, competence and fratricide in *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 79, no. 11, 4550-4558.

Trollfors, B., L. Burman, E. Dannetun, J. Llompart & R. Norrby. 1983, "Serotyping of *Streptococcus pneumoniae* strains by coagglutination and counterimmunoelectrophoresis.", *J. Clin. Microbiol.,* vol. 18, no. 4, 978-980.

Trzciński, K., C. M. Thompson & M. Lipsitch. 2004, "Single-step capsular transformation and acquisition of penicillin resistance in *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 186, no. 11, 3447-3452.

Tseng, H.-J., A. G. McEwan, J. C. Paton & M. P. Jennings. 2002, "Virulence of *Streptococcus pneumoniae*: PsaA mutants are hypersensitive to oxidative stress", *Infect. Immun.,* vol. 70, no. 3, 1635-1639.

Tu, A. H., R. L. Fulgham, M. A. McCrory, D. E. Briles & A. J. Szalai. 1999, "Pneumococcal surface protein A inhibits complement activation by *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 67, no. 9, 4720-4724.

Tuomanen, E., H. Liu, B. Hengstler, O. Zak & A. Tomasz. 1985, "The induction of meningeal inflammation by components of the pneumococcal cell wall.", *J. Infect. Dis.,* vol. 151, no. 5, 859-868.

Turner, P., J. Hinds, C. Turner, A. Jankhot, K. Gould, S. D. Bentley, F. Nosten & D. Goldblatt. 2011, "Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray", *J. Clin. Microbiol.,* vol. 49, no. 5, 1784-1789.

Uriyo, J., R. D. Gosling, V. Maddox, N. E. Sam, W. Schimana, S. H. Gillespie & T. D. McHugh. 2006, "Prevalences of *Pneumocystis jiroveci*, *Mycobacterium tuberculosis* and *Streptococcus pneumoniae* infection in children with severe pneumonia, in a tertiary referral hospital in northern Tanzania.", *Ann. Trop. Med. Parasitol.*, vol. 100, no. 3, 245-249.

Valente, C., J. Hinds, F. Pinto, S. D. Brugger, K. Gould, K. Mühlemann, H. de Lencastre & R. Sá-Leão. 2012, "Decrease in pneumococcal co-colonization following vaccination with the seven-valent pneumococcal conjugate vaccine.", *PLoS ONE*, vol. 7, no. 1, e30235.

Van Bambeke, F., R. R. Reinert, P. C. Appelbaum, P. M. Tulkens & W. E. Peetermans. 2007, "Multidrug-resistant *Streptococcus pneumoniae* infections: current and future therapeutic options.", *Drugs*, vol. 67, no. 16, 2355-2382.

van Gils, E. J. M., E. Hak, R. H. Veenhoven, G. D. Rodenburg, D. Bogaert, J. P. Bruin, L. Van Alphen & E. A. M. Sanders. 2011a, "Effect of seven-valent pneumococcal conjugate vaccine on *Staphylococcus aureus* colonisation in a randomised controlled trial.", *PLoS ONE*, vol. 6, no. 6, e20229.

van Gils, E. J. M., R. H. Veenhoven, G. D. Rodenburg, E. Hak & E. A. M. Sanders. 2011b, "Effect of 7-valent pneumococcal conjugate vaccine on nasopharyngeal carriage with *Haemophilus influenzae* and *Moraxella catarrhalis* in a randomized controlled trial", *Vaccine*, vol. 29, no. 44, 7595-7598.

van Rossum, A. M. C., E. S. Lysenko & J. N. Weiser. 2005, "Host and bacterial factors contributing to the clearance of colonization by *Streptococcus pneumoniae* in a murine model", *Infect. Immun.*, vol. 73, no. 11, 7718-7726.

van Selm, S., M. A. B. Kolkman, B. A. M. van der Zeijst, K. A. Zwaagstra, W. Gaastra & J. P. M. van Putten. 2002, "Organization and characterization of the capsule biosynthesis locus of *Streptococcus pneumoniae* serotype 9V", *Microbiology*, vol. 148, no. 6, 1747-1755.

Varvio, S.-L., K. Auranen, E. Arjas & P. H. Mäkelä. 2009, "Evolution of the capsular regulatory genes in *Streptococcus pneumoniae*", *J. Infect. Dis.*, vol. 200, no. 7, 1144-1151.

Vestrheim, D., E. Hoiby, I. Aaberge & D. Caugant. 2008a, "Phenotypic and genotypic characterization of *Streptococcus pneumoniae* strains colonizing children attending day-care centers in Norway", *J. Clin. Microbiol.*, vol. 46, no. 8, 2508-2518.

Vestrheim, D., E. Hoiby, I. Aaberge & D. Caugant. 2010, “Impact of a pneumococcal conjugate vaccination program on carriage among children in Norway”, *Clin. Vaccine Immunol.*, vol. 17, no. 3, 325-334.

Vestrheim, D. F., P. Gaustad, I. S. Aaberge & D. A. Caugant. 2011, "Pherotypes of pneumococcal strains co-existing in healthy children", *Infect. Genet. Evol.* vol. 11, no. 7, 1703-1708.

Vestrheim, D. F., E. A. Høiby, I. S. Aaberge & D. A. Caugant. 2008b, "Phenotypic and genotypic characterization of *Streptococcus pneumoniae* strains colonizing children attending day-care centers in Norway", *J. Clin. Microbiol.*, vol. 46, no. 8, 2508-2518.

Vincze, T. 2003, "NEBcutter: a program to cleave DNA with restriction enzymes", *Nucleic Acids Res.*, vol. 31, no. 13, 3688-3691.

Vives, M., M. E. Garcia, P. Saenz, M. A. Mora, L. Mata, H. Sabharwal & C. Svanborg. 1997, "Nasopharyngeal colonization in Costa Rican children during the first year of life", *Pediatr. Infect. Dis. J.*, vol. 16, no. 9, 852-858.

Vollmer, W. & A. Tomasz. 2000, "The *pgdA* gene encodes for a peptidoglycan *N*-acetylglucosamine deacetylase in Streptococcus pneumoniae.", *J. Biol. Chem.*, vol. 275, no. 27, 20496-20501.

Vollmer, W. & A. Tomasz. 2001, "Identification of the teichoic acid phosphorylcholine esterase in *Streptococcus pneumoniae*", *Mol. Microbiol.*, vol. 39, no. 6, 1610-1622.

Waite, R. D., D. W. Penfold, J. K. Struthers & C. G. Dowson. 2003, "Spontaneous sequence duplications within capsule genes *cap8E* and *tts* control phase variation in *Streptococcus pneumoniae* serotypes 8 and 37", *Microbiology*, vol. 149, no. 2, 497-504.

Waite, R. D., J. K. Struthers & C. G. Dowson. 2001, "Spontaneous sequence duplication within an open reading frame of the pneumococcal type 3 capsule locus causes high-frequency phase variation", *Mol. Microbiol.*, vol. 42, no. 5, 1223-1232.

Walsh, P. S., D. A. Metzger & R. Higuchi. 1991, "Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material", *Biotechniques*, vol. 10, no. 4, 506-513.

Wang, Q., M. Wang, F. Kong, G. L. Gilbert, B. Cao, L. Wang & L. Feng. 2007, "Development of a DNA microarray to identify the *Streptococcus pneumoniae* serotypes contained in the 23-valent pneumococcal polysaccharide vaccine and closely related serotypes.", *J. Microbiol. Methods,* vol. 68, no. 1, 128-136.

Wani, J. H., J. V. Gilbert, A. G. Plaut & J. N. Weiser. 1996, "Identification, cloning, and sequencing of the immunoglobulin A1 protease gene of *Streptococcus pneumoniae*", *Infect. Immun.,* vol. 64, no. 10, 3967-3974.

Wasas, A. D., R. E. Huebner, M. De Blanche & K. P. Klugman. 1998, "Long-term survival of *Streptococcus pneumoniae* at room temperature on Dorset egg medium", *J. Clin. Microbiol.,* vol. 36, no. 4, 1139-1140.

Watson, D. A. & D. M. Musher. 1990, "Interruption of capsule production in *Streptococcus pneumonia* serotype 3 by insertion of transposon Tn916", *Infect. Immun.,* vol. 58, no. 9, 3135-3138.

Watson, D. A. & D. M. Musher. 1999, "A brief history of the pneumococcus in biomedical research", *Semin. Respir. Infect.,* vol. 14, no. 3, 198-208.

Watson, D. A., D. M. Musher, J. W. Jacobson & J. Verhoef. 1993, "A brief history of the pneumococcus in biomedical research: A panoply of scientific discovery", *Clin. Infect. Dis.,* vol. 17, no. 5, 913-924.

Watt, J. P., K. L. O'Brien, S. Katz, M. A. Bronsdon, J. Elliott, J. Dallas, M. J. Perilla, R. Reid, L. Murrow, R. Facklam, M. Santosham & C. G. Whitney. 2004, "Nasopharyngeal versus oropharyngeal sampling for detection of pneumococcal carriage in adults", *J. Clin. Microbiol.,* vol. 42, no. 11, 4974-4976.

Ween, O., P. Gaustad & L. S. Håvarstein. 1999, "Identification of DNA binding sites for ComE, a key regulator of natural competence in *Streptococcus pneumoniae*", *Mol. Microbiol.,* vol. 33, no. 4, 817-827.

Weimer, K. E. D., C. E. Armbruster, R. A. Juneau, W. Hong, B. Pang & W. E. Swords. 2010, "Coinfection with *Haemophilus influenzae* promotes pneumococcal biofilm formation during experimental otitis media and impedes the progression of pneumococcal disease", *J. Infect. Dis.,* vol. 202, no. 7, 1068-1075.

Weinberger, D. M., R. Dagan, N. Givon-Lavi, G. Regev-Yochay, R. Malley & M. Lipsitch. 2008, "Epidemiologic evidence for serotype-specific acquired

immunity to pneumococcal carriage", *J. Infect. Dis.*, vol. 197, no. 11, 1511-1518.

Weinberger, D. M., K. Trzciński, Y.-J. Lu, D. Bogaert, A. Brandes, J. Galagan, P. W. Anderson, R. Malley & M. Lipsitch. 2009, "Pneumococcal capsular polysaccharide structure predicts serotype prevalence", *PLoS Pathog.*, vol. 5, no. 6, e1000476.

Weiser, J. N., R. Austrian, P. K. Sreenivasan & H. R. Masure. 1994, "Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization", *Infect. Immun.*, vol. 62, no. 6, 2582-2589.

Weiser, J. N., D. Bae, H. Epino, S. B. Gordon, M. Kapoor, L. A. Zenewicz & M. Shchepetov. 2001, "Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 69, no. 9, 5430-5439.

Weiser, J. N., D. Bae, C. Fasching, R. W. Scamurra, A. J. Ratner & E. N. Janoff. 2003, "Antibody-enhanced pneumococcal adherence requires IgA1 protease", *Proc. Nat. Acad. Sci. U. S. A.*, vol. 100, no. 7, 4215-4220.

Weiser, J. N., J. B. Goldberg, N. Pan, L. Wilson & M. Virji. 1998a, "The phosphorylcholine epitope undergoes phase variation on a 43-kilodalton protein in *Pseudomonas aeruginosa* and on pili of *Neisseria meningitidis* and *Neisseria gonorrhoeae*", *Infect. Immun.*, vol. 66, no. 9, 4263-4267.

Weiser, J. N. & M. Kapoor. 1999, "Effect of intrastrain variation in the amount of capsular polysaccharide on genetic transformation of *Streptococcus pneumoniae*: implications for virulence studies of encapsulated strains.", *Infect. Immun.*, vol. 67, no. 7, 3690-3692.

Weiser, J. N., N. Pan, K. L. McGowan, D. Musher, A. Martin & J. Richards. 1998b, "Phosphorylcholine on the lipopolysaccharide of *Haemophilus influenzae* contributes to persistence in the respiratory tract and sensitivity to serum killing mediated by C-reactive protein", *J. Exp. Med.*, vol. 187, no. 4, 631-640.

Wellmer, A., G. Zysk, J. Gerber, T. Kunst, M. Von Mering, S. Bunkowski, H. Eiffert & R. Nau. 2002, "Decreased virulence of a pneumolysin-deficient strain of *Streptococcus pneumoniae* in murine meningitis", *Infect. Immun.*, vol. 70, no. 11, 6504-6508.

Werno, A. M. & D. R. Murdoch. 2008, "Medical microbiology: laboratory diagnosis of invasive pneumococcal disease", *Clin. Infect. Dis.,* vol. 46, no. 6, 926-932.

Wester, C. W., D. Ariga, C. Nathan, T. W. Rice, J. Pulvirenti, R. Patel, F. Kocka, J. Ortiz & R. A. Weinstein. 2002, "Possible overestimation of penicillin resistant *Streptococcus pneumoniae* colonization rates due to misidentification of oropharyngeal streptococci", *Diagn. Microbiol. Infect. Dis.,* vol. 42, no. 4, 263-268.

Whalan, R. H., S. G. P. Funnell, L. D. Bowler, M. J. Hudson, A. Robinson & C. G. Dowson. 2005, "PiuA and PiaA, iron uptake lipoproteins of *Streptococcus pneumoniae*, elicit serotype independent antibody responses following human pneumococcal septicaemia", *FEMS Immunol. Med. Microbiol.,* vol. 43, no. 1, 73-80.

Whalan, R. H., S. G. P. Funnell, L. D. Bowler, M. J. Hudson, A. Robinson & C. G. Dowson. 2006, "Distribution and genetic diversity of the ABC transporter lipoproteins PiuA and PiaA within *Streptococcus pneumoniae* and related streptococci", *J. Bacteriol.,* vol. 188, no. 3, 1031-1038.

Whatmore, A., S. King, N. Doherty, D. Sturgeon, N. Chanter & C. Dowson. 1999, "Molecular characterization of equine isolates of *Streptococcus pneumoniae*: natural disruption of genes encoding the virulence factors pneumolysin and autolysin", *Infect. Immun.,* vol. 67, no. 6, 2776-2782.

Whatmore, A. M., V. A. Barcus & C. G. Dowson. 1999, "Genetic diversity of the streptococcal competence (*com*) gene locus", *J. Bacteriol.,* vol. 181, no. 10, 3144-3154.

Whatmore, A. M., A. Efstratiou, A. P. Pickerill, K. Broughton, G. Woodard, D. Sturgeon, R. George & C. G. Dowson. 2000, "Genetic relationships between clinical isolates of *Streptococcus pneumoniae*, *Streptococcus oralis*, and *Streptococcus mitis*: characterization of 'atypical' pneumococci and organisms allied to *S. mitis* harboring *S. pneumoniae* virulence factor-encoding genes", *Infect. Immun.,* vol. 68, no. 3, 1374-1382.

Whitfield, C. & A. Paiment. 2003, "Biosynthesis and assembly of Group 1 capsular polysaccharides in *Escherichia coli* and related extracellular polysaccharides in other bacteria", *Carbohydr. Res.,* vol. 338, no. 23, 2491-2502.

Whitney, C. G., M. M. Farley, J. Hadler, L. H. Harrison, N. M. Bennett, R. Lynfield, A. Reingold, P. R. Cieslak, T. Pilishvili, D. Jackson, R. R. Facklam, J. H. Jorgensen, A. Schuchat & A. B. C. S. o. t. E. I. P. Network. 2003, "Decline in invasive pneumococcal disease after the introduction of protein-

polysaccharide conjugate vaccine.", *N. Engl. J. Med.*, vol. 348, no. 18, 1737-1746.

Wicher, K., C. Kalinka, P. Mlodozeniec & N. R. Rose. 1982, "Fluorescent antibody technic used for identification and typing of *Streptococcus pneumoniae*", *Am. J. Clin. Pathol.*, vol. 77, no. 1, 72-77.

Wiktor, S. Z., M. Sassan-Morokro, A. D. Grant, L. Abouya, J. M. Karon, C. Maurice, G. Djomand, A. Ackah, K. Domoua, A. Kadio, A. Yapi, P. Combe, O. Tossou, T. H. Roels, E. M. Lackritz, D. Coulibaly, K. M. De Cock, I. M. Coulibaly & A. E. Greenberg. 1999, "Efficacy of trimethoprim-sulphamethoxazole prophylaxis to decrease morbidity and mortality in HIV-1-infected patients with tuberculosis in Abidjan, Côte d'Ivoire: a randomised controlled trial", *Lancet*, vol. 353, no. 9163, 1469-1475.

Wilén, M., W. Buwembo, H. Sendagire, F. Kironde & G. Swedberg. 2009, "Cotrimoxazole resistance of *Streptococcus pneumoniae* and commensal streptococci from Kampala, Uganda", *Scand. J. Infect. Dis.*, vol. 41, no. 2, 113-121.

Woo, P. C., E. Y. Cheung, K. Leung & K. Yuen. 2001, "Identification by 16S ribosomal RNA gene sequencing of an *Enterobacteriaceae* species with ambiguous biochemical profile from a renal transplant recipient", *Diagn. Microbiol. Infect. Dis.*, vol. 39, no. 2, 85-93.

World Health Organization. 2007, "World Health Organization: Pneumococcal conjugate vaccine for childhood immunization-WHO position paper.", *Wkly. Epidemiol. Rec.*, vol. 82, no. 93-104.

World Health Organization. (2008). "Morbidity and Mortality Weekly Report: Progress in Introduction of Pneumococcal Conjugate Vaccine - Worldwide, 2000-2008." 2011, from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5742a2.htm.

Wroe, P. C., G. M. Lee, J. A. Finkelstein, S. I. Pelton, W. P. Hanage, M. Lipsitch, A. E. Stevenson, S. L. Rifas-Shiman, K. Kleinman, M. M. Dutta-Linn, V. L. Hinrichsen, M. Lakoma & S. S. Huang. 2011, "Pneumococcal carriage and antibiotic resistance in young children before 13-valent conjugate vaccine.", *Pediatr. Infect. Dis. J.*, vol. 31, no. 3, 249-254.

Wu, H., M. Nahm & Y. Guo. 1997, "Intranasal immunization of mice with PspA (pneumococcal surface protein A) can prevent intranasal carriage, pulmonary infection, and sepsis with *Streptococcus pneumoniae*", *J. Infect. Dis.*, vol. 175, no. 4, 839-846.

Xu, Q., M. E. Pichichero, J. R. Casey & M. Zeng. 2009, "Novel type of *Streptococcus pneumoniae* causing multidrug-resistant acute otitis media in children", *Emerg. Infect. Dis.,* vol. 15, no. 4, 547-551.

Yagupsky, P., N. Porat, D. Fraser, F. Prajgrod, M. Merires, L. Mcgee, K. P. Klugman & R. Dagan. 1998, "Acquisition, carriage, and transmission of pneumococci with decreased antibiotic susceptibility in young children attending a day care facility in southern Israel", *J. Infect. Dis.,* vol. 177, no. 4, 1003-1012.

Yamamoto, M., L. S. McDaniel, K. Kawabata, D. E. Briles, R. J. Jackson, J. R. McGhee & H. Kiyono. 1997, "Oral immunization with PspA elicits protective humoral immunity against *Streptococcus pneumoniae* infection.", *Infect. Immun.,* vol. 65, no. 2, 640-644.

Yaro, S., M. Lourd, Y. Traore, B.-M. Njanpop Lafourcade, A. Sawadogo, L. Sangare, A. Hien, M. S. Ouedraogo, O. Sanou, I. Parent du Châtelet, J.-L. Koeck & B. D. Gessner. 2006, "Epidemiological and molecular characteristics of a highly lethal pneumococcal meningitis epidemic in Burkina Faso.", *Clin. Infect. Dis.,* vol. 43, no. 6, 693-700.

Yeowell, H. N. & J. R. White. 1982, "Iron requirement in the bactericidal mechanism of streptonigrin", *Antimicrob. Agents Chemother.,* vol. 22, no. 6, 961-968.

Yother, J. (2004). Capsules. The Pneumococcus. E. Tuomanen, T. J. Mitchell, D. A. Morrison and B. Spratt. Washington, D.C., ASM Press**:** 30-48.

Yother, J. 2011, "Capsules of *Streptococcus pneumoniae* and other bacteria: paradigms for polysaccharide biosynthesis and regulation", *Ann. Rev. Microbiol,* vol. 65, no. 563-581.

Yother, J. & D. E. Briles. 1992, "Structural properties and evolutionary relationships of PspA, a surface protein of *Streptococcus pneumoniae*, as revealed by sequence analysis", *J. Bacteriol.,* vol. 174, no. 2, 601-609.

Yother, J., L. S. McDaniel & D. E. Briles. 1986, "Transformation of encapsulated *Streptococcus pneumoniae*", *J. Bacteriol.,* vol. 168, no. 3, 1463-1465.

Yu, J., M. d. G. S. Carvalho, B. Beall & M. H. Nahm. 2008, "A rapid pneumococcal serotyping system based on monoclonal antibodies and PCR", *J. Med. Microbiol.,* vol. 57, no. 2, 171-178.

Yu, J., J. Lin, W. H. Benjamin, K. B. Waites, C.-h. Lee & M. H. Nahm. 2005, "Rapid multiplex assay for serotyping pneumococci with monoclonal and polyclonal antibodies", *J. Clin. Microbiol*., vol. 43, no. 1, 156-162.

Yu, J., J. Lin, K. H. Kim, W. H. Benjamin & M. H. Nahm. 2011, "Development of an automated and multiplexed serotyping assay for *Streptococcus pneumoniae*", *Clin. Vaccine Immun.*, vol. 18, no. 11, 1900-1907.

Yu, V. L., C. C. C. Chiou, C. Feldman, A. Ortqvist, J. Rello, A. J. Morris, L. M. Baddour, C. M. Luna, D. R. Snydman, M. Ip, W. C. Ko, M. B. F. Chedid, A. Andremont, K. P. Klugman & I. P. S. Group. 2003, "An international prospective study of pneumococcal bacteremia: correlation with in vitro resistance, antibiotics administered, and clinical outcome", *Clin. Infect. Dis.*, vol. 37, no. 2, 230-237.

Zähner, D., A. R. Gandhi, H. Yi & D. S. Stephens. 2011, "Mitis group streptococci express variable pilus islet 2 pili.", *PLoS ONE,* vol. 6, no. 9, e25124.

Zhang, A., M. Yang, P. Hu, J. Wu, B. Chen, Y. Hua, J. Yu, H. Chen, J. Xiao & M. Jin. 2011, "Comparative genomic analysis of *Streptococcus suis* reveals significant genomic diversity among different serotypes.", *BMC Genomics,* vol. 12, no. 523.

Zhang, J. R., K. E. Mostov, M. E. Lamm, M. Nanno, S. Shimida, M. Ohwaki & E. Tuomanen. 2000, "The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells", *Cell,* vol. 102, no. 6, 827-837.

Zhang, W., G. Y. Hu & A. Deisseroth. 1991, "Improvement of PCR sequencing by formamide", *Nucleic Acids Res.,* vol. 19, no. 23, 6649.

Zhang, Y., K. Auranen & M. Eichner. 2004, "The influence of competition and vaccination on the coexistence of two pneumococcal serotypes.", *Epidemiol. Infect.*, vol. 132, no. 6, 1073-1081.

Zhou, F., F. Kong, Z. Tong & G. L. Gilbert. 2007, "Identification of less-common *Streptococcus pneumoniae* serotypes by a multiplex PCR-based reverse line blot hybridization assay", *J. Clin. Microbiol.*, vol. 45, no. 10, 3411-3415.

Zhu, L., Y. Zhang, J. Fan, M. C. Herzberg & J. Kreth. 2011, "Characterization of competence and biofilm development of a *Streptococcus sanguinis* endocarditis isolate", *Mol. Oral Microbiol.,* vol. 26, no. 2, 117-126.

Zysk, G., B. K. Schneider-Wald, J. H. Hwang, L. Bejo, K. S. Kim, T. J. Mitchell, R. Hakenbeck & H. P. Heinz. 2001, "Pneumolysin is the main inducer of cytotoxicity to brain microvascular endothelial cells caused by *Streptococcus pneumoniae*", *Infect. Immun.*, vol. 69, no. 2, 845-852.

Appendix Table A1. Strains used in the works of this thesis

Strain Collection	Strain Name	Serotype by Conventional Typing	Sequetyping with GenBank database	Serotype-specific DNA-PCR[a]	*comC* allele (pherotype)	Allelic Profile (ST)	MIC (µg/mL)[b]		Chapter 3	Chapter 4	Chapter 5	Chapter 6
							Pen	Sxt				
Tanzania (Multiple Colonization Study)	1NP10S	NT	no PCR	NP	4	4156	0.032	0.19		√		
	1NP7	NT	no PCR	NP	4	4156	0.032	0.19		√		
	1NP4	NT	no PCR	NP	4	4156	0.032	0.19		√		
	1NP5	NT	no PCR	NP	4	4156	0.032	0.19		√		
	1NP9	NT	no PCR	NP	4	4156	0.032	0.19		√		
	1NP10B	6B	6B	NP	2.1	4429	0.19	6		√	√	
	11NP8 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11OP5 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11OP9 (Jan/Feb)	6B	6B	NP	NP	4432	0.25	6		√		
	11NP5 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11OP6 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	8		√		
	11OP10 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP2 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP3 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP4 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP6 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP7 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP9 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11NP10 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6	√	√		√
	11NP1 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	11OP3 (Jan/Feb)	6B	6B	NP	NP	4432	0.19	6		√		
	16OP4	19F	19F	NP	1	347	0.19	6	√	√		√
	16NP1	19F	19F	NP	1	347	0.19	6		√		
	16NP4	19F	19F	NP	1	347	0.19	6		√	√	
	16OP5	19F	19F	NP	1	347	0.19	6		√		
	16OP2	19F	19F	NP	1	347	0.19	6		√		
	16NP3	19F	19F	NP	1	347	0.19	6		√		
	16NP5	34	17A/34	34	2.1	4158	0.032	6		√	√	
	16NP8	34	17A/34	34	2.1	4158	0.032	6		√		
	16NP9	34	17A/34	34	2.1	4158	0.032	6		√		
	16NP7	34	17A/34	34	2.1	4158	0.032	6		√		
	16NP10	35A	35B/C	NP	4	840	0.25	4	√	√	√	√
	18NP1	10A	10A	NP	1	852	0.125	4		√	√	
	18NP5	10A	10A	NP	1	852	0.125	4		√		

	18NP9	10A	10A	NP	1	852	0.125	4		√			
	18NP10	10A	10A	NP	1	852	0.125	4		√			
	22NP4	17F	17F	NP	2.1	4160	0.016	0.125		√	√		
	22NP7	17F	17F	NP	2.1	4160	0.016	0.125		√			
	22NP9	17F	17F	NP	2.1	4160	0.016	0.125		√			
	22NP10	17F	17F	NP	2.1	4160	0.016	0.125		√			
	22NP2	17F	17F	NP	2.1	4160	0.016	0.125	√	√		√	
	35NP4	10A	10A	NP	1	852	0.38	6		√	√		
	35NP10	19A	19A	NP	2.1	4162	0.125	2	√	√	√	√	√
	35NP9	10A	10A	NP	1	852	0.38	6		√			
	35NP1	6B	6B	NP	4	4157	0.125	6	√	√	√	√	
	35NP2	21/39	21	NP	1	1145	0.064	3		√			
	35NP6	6B	6B	NP	2.1	854	0.125	6	√	√	√	√	
	35NP8	19F	19F	NP	1	6170	0.19	2	√	√	√	√	
	45OP1	21/39	21	NP	1	1145	0.094	4		√	√		
	45OP2	21/39	21	NP	1	1145	0.094	4		√			
	45OP3	21/39	21	NP	1	1145	0.094	4		√			
	45OP4	21/39	21	NP	1	1145	0.094	4		√			
	45OP5	21/39	21	NP	1	1145	0.094	4		√			
	45OP6	21/39	21	NP	1	1145	0.094	4		√			
	45OP7	21/39	21	NP	1	1145	0.094	4		√			
	45OP8	21/39	21	NP	1	1145	0.094	4		√			
	45OP9	21/39	21	NP	1	1145	0.094	4	√	√		√	
	55OP3	35A	35B/C	NP	NP	NP	0.19	3	√	√		√	
	69OP8	6B	6B	NP	NP	4429	0.19	8		√			
	69OP9	6B	6B	NP	NP	4429	0.19	8		√			
	69OP4	6B	6B	NP	NP	4429	0.19	8		√			
	69OP3	6B	6B	NP	NP	4429	0.19	8		√			
	69OP7	6B	6B	NP	NP	4429	0.19	8		√			
	69OP2	6B	6B	NP	NP	4429	0.19	8		√			
	69OP10	6B	6B	NP	NP	4429	0.19	8	√	√		√	
	3NP9 (Mar)	6B	6B	NP	4	4368	0.032	0.125		√	√		
	3NP6 (Mar)	6B	6B	NP	4	4368	0.032	0.19		√			
	3NP7 (Mar)	6B	6B	NP	4	4368	0.047	0.125	√	√		√	√
	10NP1 (Mar)	13/28	13/20	13	4	4370	0.125	6		√	√		
	10NP4 (Mar)	13/28	13/20	13	4	4370	0.125	6		√			
	10NP5 (Mar)	13/28	13/20	13	4	4370	0.125	6		√			
	10NP7 (Mar)	13/28	13/20	13	4	4370	0.125	6		√			
	10NP8 (Mar)	13/28	13/20	13	4	4370	0.125	6		√			

10NP10 (Mar)	13/28	13/20	13	4	4370	0.125	6		√		
11OP1 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP2 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP3 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP4 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP5 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP6 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP8 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
11OP9 (Jun)	6B	6B	NP	NP	4432	0.19	4		√		
3OP1 (Sep)	1	1	NP	1	217	<0.016	<0.002	√	√	√	√
3OP2 (Sep)	1	1	NP	1	217	<0.016	<0.002		√		
3OP3 (Sep)	1	1	NP	1	217	<0.016	<0.002		√		
3OP6 (Sep)	1	1	NP	1	217	<0.016	<0.002		√		
3OP7 (Sep)	6B	6B	NP	4	4373	0.25	6	√	√	√	√
3OP8 (Sep)	6B	6B	NP	4	4373	0.25	6		√		
3OP9 (Sep)	6B	6B	NP	4	4373	0.25	6		√		
3OP10 (Sep)	1	1	NP	1	217	S	S		√		
21NP1	10A	10A	NP	NP	NP	2	1.5	√	√		√
24OP1	19A	19A	NP	2.1	4162	0.25	6		√	√	
24OP3	19A	19A	NP	2.1	4162	0.25	6		√	√	
24OP5	19A	19A	NP	2.1	4162	0.25	6		√		
29OP8	35A	35B/35C	NP	4	840	0.25	4		√		
29OP10	11A	11A/D/18F	NP	4	5752	0.38	6	√	√	√	√
3NP8 (Jan)	6B	6B	NP	4	4368	S	S		√	√	
3OP1 (Jul)	1	1	NP	1	217	S	S		√	√	
3OP2 (Jul)	1	1	NP	1	217	S	S		√		
3OP4 (Jul)	1	1	NP	1	217	S	S		√		
3OP5 (Jul)	1	1	NP	1	217	S	S		√		
3OP6 (Jul)	1	1	NP	1	217	S	S		√		
7NP4	21/39	21	NP	NP	NP	0.38	8	√	√		√
7OP9	23B	23B	NP	NP	NP	S	1.5	√	√		√
10NP2 (Jan/Feb)	19A	19A	NP	2.1	4162	0.125	6		√	√	
10NP3 (Jan/Feb)	19A	19A	NP	2.1	4162	0.125	6		√		
10NP5 (Jan/Feb)	19A	19A	NP	2.1	4162	0.125	6		√		
10NP6 (Jan/Feb)	19A	19A	NP	2.1	4162	0.125	6		√		
10OP6 (Jan/Feb)	13/28	13/20	13	4	4370	0.125	3		√	√	
10NP7 (Jan/Feb)	19A	19A	NP	2.1	4162	0.125	6		√	√	
10NP9 (Jan/Feb)	19A	19A	NP	2.1	4162	0.125	6		√		
10OP1 (Apr)	18C	18B/C	NP	1	NP	S	S		√		

	10OP3 (Apr)	18C	18B/C	NP	1	NP	S	S		√	√	
	10OP5 (Apr)	18C	18B/C	NP	1	NP	S	S		√		
	10OP6 (Apr)	18C	18B/C	NP	1	NP	S	S		√		
	10OP7 (Apr)	18C	18B/C	NP	1	NP	S	S		√		
	10OP9 (Apr)	6B	6B	NP	4	NP	S	S		√	√	
	10OP10 (Apr)	18C	18B/C	NP	1	NP	S	S		√		
	11NP1 (May)	4	4	NP	NP	NP	S	S		√		
	11NP2 (May)	4	4	NP	NP	NP	S	S		√		
	11NP3 (May)	4	4	NP	NP	NP	S	S		√		
	11NP4 (May)	4	4	NP	NP	NP	S	S		√		
	11NP5 (May)	4	4	NP	NP	NP	S	S		√		
	11NP6 (May)	4	4	NP	NP	NP	S	S		√		
	11NP7 (May)	4	4	NP	NP	NP	S	S		√		
	11NP8 (May)	4	4	NP	NP	NP	S	S		√		
Tanzania (Trachoma Study)	1003-2	7C	7C	NP	NP	NP	NP	NP	√			√
	2514-1	7C	7C	NP	NP	NP	NP	NP	√			√
	1804.1	NP	NP	NP	1	NP	NP	NP			√	
	3205.2	NP	NP	NP	1	NP	NP	NP			√	
	3207.3	NP	NP	NP	1	NP	NP	NP			√	
	4407.1	NP	NP	NP	1	NP	NP	NP			√	
	5411.1	NP	NP	NP	1	NP	NP	NP			√	
	5711.1	NP	NP	NP	1	NP	NP	NP			√	
	5802.1	NP	NP	NP	1	NP	NP	NP			√	
	5907.2	NP	NP	NP	1	NP	NP	NP			√	
	6611.1	NP	NP	NP	1	NP	NP	NP			√	
	6901.2	NP	NP	NP	1	NP	NP	NP			√	
	6903.2	NP	NP	NP	2.1	NP	NP	NP			√	
	107.1	NP	NP	NP	4	NP	NP	NP			√	
	112.1	NP	NP	NP	1	NP	NP	NP			√	
	210.3	NP	NP	NP	2.1	NP	NP	NP			√	
	505.2	NP	NP	NP	1	NP	NP	NP			√	
	602.1	NP	NP	NP	2.1	NP	NP	NP			√	
	607.3	NP	NP	NP	1	NP	NP	NP			√	
	2711.3	NP	NP	NP	4	NP	NP	NP			√	
	2801.2	NP	NP	NP	4	NP	NP	NP			√	
	2810.1	NP	NP	NP	4	NP	NP	NP			√	
	2907.1	NP	NP	NP	2.1	NP	NP	NP			√	

	3612.2	NP	NP	NP	2.1	NP	NP	NP			√	
	3612.3	NP	NP	NP	4	NP	NP	NP			√	
	4407.1	NP	NP	NP	1	NP	NP	NP			√	
	4806.1	NP	NP	NP	2.1	NP	NP	NP			√	
	5303.2	NP	NP	NP	2.1	NP	NP	NP			√	
	5605.1	NP	NP	NP	2.1	NP	NP	NP			√	
	5609.1	NP	NP	NP	1	NP	NP	NP			√	
	5609.2	NP	NP	NP	2.1	NP	NP	NP			√	
	5704.2	NP	NP	NP	2.1	NP	NP	NP			√	
	5711.1	NP	NP	NP	1	NP	NP	NP			√	
	5803.1	NP	NP	NP	2.1	NP	NP	NP			√	
	5808.1	NP	NP	NP	1	NP	NP	NP			√	
	6614.2	NP	NP	NP	4	NP	NP	NP			√	
	6802.2	NP	NP	NP	4	NP	NP	NP			√	
	7502.2	NP	NP	NP	4	NP	NP	NP			√	
	8715.2	NP	NP	NP	2.1	NP	NP	NP			√	
	8715.4	NP	NP	NP	2.1	NP	NP	NP			√	
	8802.2	NP	NP	NP	4	NP	NP	NP			√	
Tanzania Invasive Isolate (John Crump)	JC001	NP	NP	NP	1	NP	0.125	9			√	
	JC002	NP	NP	NP	1	NP	0.016	20			√	
	JC003	NP	NP	NP	3	NP	0.016	12			√	
	JC004	NP	NP	NP	1	NP	0.016	11			√	
	JC005	NP	NP	NP	1	NP	0.023	20			√	
	JC006	NP	NP	NP	1	NP	0.25	11			√	
	JC007	NP	NP	NP	2.1	NP	0.25	6			√	
	JC008	NP	NP	NP	1	NP	0.016	19			√	
	JC009	NP	NP	NP	2.1	NP	0.25	15			√	
	JC010	NP	NP	NP	2.1	NP	0.19	10			√	
	JC011	NP	NP	NP	4	NP	0.125	13			√	
	JC012	NP	NP	NP	2.1	NP	0.25	17			√	
	JC013	NP	NP	NP	2.1	NP	0.047	23			√	
	JC014	NP	NP	NP	4	NP	0.25	10			√	
	JC015	NP	NP	NP	1	NP	0.016	20			√	
	JC016	NP	NP	NP	1	NP	0.016	5			√	
	JC017	NP	NP	NP	4	NP	0.25	7			√	

	JC018	NP	NP	NP	1	NP	0.023	20			√	
United Kingdom Royal Free Hospital Strains	N3	NP	NP	NP	1	NP	NP	NP			√	
	N5	NP	NP	NP	1	NP	NP	NP			√	
	N6	NP	NP	NP	1	NP	NP	NP			√	
	N7	NP	NP	NP	1	NP	NP	NP			√	
	N8	NP	NP	NP	1	NP	NP	NP			√	
	N9	NP	NP	NP	1	NP	NP	NP			√	
	N10	NP	NP	NP	1	NP	NP	NP			√	
	N13	NP	NP	NP	2.2	NP	NP	NP			√	
	N15	6	6A	6A	2.1	65	NP	NP	√		√	√
	N16	NP	NP	NP	1	NP	NP	NP			√	
	N18	NP	NP	NP	1	NP	NP	NP			√	
	N19	NP	NP	NP	4	NP	NP	NP			√	
	N21	NP	NP	6A	2.1	65	NP	NP	√		√	
	N22	NP	NP	NP	1	NP	NP	NP			√	
	N27	NP	NP	NP	1	NP	NP	NP			√	
	N32	NP	NP	NP	1	NP	NP	NP			√	
	N34	NP	NP	NP	1	NP	NP	NP			√	
	N36	NP	NP	NP	1	NP	NP	NP			√	
	N38	NP	NP	NP	1	NP	NP	NP			√	
	N39	NP	NP	NP	1	NP	NP	NP			√	
	N40	NP	NP	NP	1	NP	NP	NP			√	
	N41	NP	NP	NP	1	NP	NP	NP			√	
	N42	NP	NP	NP	2.1	NP	NP	NP			√	
	N47	NP	NP	NP	1	NP	NP	NP			√	
	N48	NP	NP	NP	1	NP	NP	NP			√	
	N53	NP	NP	NP	2.1	NP	NP	NP			√	
	N54	NP	NP	NP	2.1	NP	NP	NP			√	
	N59	NP	NP	NP	1	NP	NP	NP			√	
	N60	NP	NP	NP	1	NP	NP	NP			√	
	N63	NP	NP	NP	1	NP	NP	NP			√	
	N64	NP	NP	NP	2.2	NP	NP	NP			√	
	N66	NP	NP	NP	1	NP	NP	NP			√	
	N70	NP	NP	NP	1	NP	NP	NP			√	
	N73	NP	NP	NP	1	NP	NP	NP			√	

N76	NP	NP	NP	1	NP	NP	NP			√	
N78	NP	NP	NP	2.1	NP	NP	NP			√	
N79	NP	NP	NP	1	NP	NP	NP			√	
N81	NP	NP	NP	1	NP	NP	NP			√	
N92	NP	NP	NP	1	NP	NP	NP			√	
N93	NP	NP	NP	1	NP	NP	NP			√	
N94	6	6A	6A	2.1	65	NP	NP	√		√	√
N98	NP	NP	NP	1	NP	NP	NP			√	
N100	NP	NP	NP	1	NP	NP	NP			√	
N105	NP	NP	NP	1	NP	NP	NP			√	
N106	NP	NP	NP	1	NP	NP	NP			√	
N107	NP	NP	NP	1	NP	NP	NP			√	
N123	NP	NP	NP	1	NP	NP	NP			√	
N125	NP	NP	NP	1	NP	NP	NP			√	
N126	NP	NP	NP	1	NP	NP	NP			√	
N128	NP	NP	NP	1	NP	NP	NP			√	
N140	NP	NP	NP	1	NP	NP	NP			√	
N147	NP	NP	NP	1	NP	NP	NP			√	
N150	NP	NP	NP	1	NP	NP	NP			√	
N151	NP	NP	NP	1	NP	NP	NP			√	
N152	NP	NP	NP	1	NP	NP	NP			√	
N155	6	6A	6A	2.1	65	NP	NP	√		√	√
N158	NP	NP	NP	1	NP	NP	NP			√	
N178	NP	NP	NP	1	NP	NP	NP			√	
N180	NP	NP	NP	1	NP	NP	NP			√	
N182	NP	NP	NP	1	NP	NP	NP			√	
N198	NP	NP	NP	1	NP	NP	NP			√	
N199	NP	NP	NP	1	NP	NP	NP			√	
N201	NP	NP	NP	1	NP	NP	NP			√	
N209	NP	NP	NP	1	NP	NP	NP			√	
N216	NP	NP	NP	1	NP	NP	NP			√	
N217	NP	NP	NP	1	NP	NP	NP			√	
N223	NP	NP	NP	2.1	NP	NP	NP			√	
N232	NP	NP	NP	1	NP	NP	NP			√	
N244	NP	NP	NP	1	NP	NP	NP			√	
N252	NP	NP	NP	1	NP	NP	NP			√	

N253	NP	NP	NP	1	NP	NP	NP			√	
N258	NP	NP	NP	1	NP	NP	NP			√	
N259	6	6A	6A	2.1	65	NP	NP	√		√	√
N362	NP	NP	NP	1	NP	NP	NP			√	
N367	NP	NP	NP	1	NP	NP	NP			√	
N369	NP	NP	NP	1	NP	NP	NP			√	
N371	NP	NP	NP	2.2	NP	NP	NP			√	
N375	NP	NP	NP	1	NP	NP	NP			√	
N380	NP	NP	NP	1	NP	NP	NP			√	
N381	NP	NP	NP	2.2	NP	NP	NP			√	
N384	NP	NP	NP	1	NP	NP	NP			√	
N388	NP	NP	NP	1	NP	NP	NP			√	
N389	NP	NP	NP	2.2	NP	NP	NP			√	
N390	NP	NP	NP	1	NP	NP	NP			√	
N391	NP	NP	NP	1	NP	NP	NP			√	
N396	NP	NP	6A	2.1	65	NP	NP			√	
N400	NP	NP	NP	1	NP	NP	NP			√	
N404	NP	NP	NP	1	NP	NP	NP			√	
N405	6	6A	6A	2.1	65	NP	NP	√		√	√
N408	NP	NP	NP	1	NP	NP	NP			√	
N447	NP	NP	NP	2.1	NP	NP	NP			√	
N448	NP	NP	NP	1	NP	NP	NP			√	
N449	NP	NP	NP	1	NP	NP	NP			√	
N450	NP	NP	NP	4	NP	NP	NP			√	
N451	NP	NP	NP	2.1	NP	NP	NP			√	
N453	NP	NP	NP	1	NP	NP	NP			√	
N454	NP	NP	NP	1	NP	NP	NP			√	
N455	NP	NP	NP	1	NP	NP	NP			√	
N456	NP	NP	NP	1	NP	NP	NP			√	
N457	NP	NP	NP	1	NP	NP	NP			√	
N458	NP	NP	NP	1	NP	NP	NP			√	
N459	NP	NP	NP	1	NP	NP	NP			√	
N460	NP	NP	NP	1	NP	NP	NP			√	
N465	NP	NP	NP	2.1	NP	NP	NP			√	
N468	NP	NP	NP	4	NP	NP	NP			√	
N2526	NP	NP	NP	1	NP	NP	NP			√	

	N2552	NP	NP	NP	1	NP	NP	NP			√	
	N5510	NP	NP	NP	1	NP	NP	NP			√	
	N21102	NP	NP	NP	1	NP	NP	NP			√	
	N161236	NP	NP	NP	1	NP	NP	NP			√	
	N273110	NP	NP	NP	1	NP	NP	NP			√	
	N888b	NP	NP	NP	1	NP	NP	NP			√	
	10M204493	NP	NP	NP	4	NP	NP	NP			√	
	10M265301	NP	NP	NP	4	NP	NP	NP			√	
	10M251479	NP	NP	NP	4	NP	NP	NP			√	
	ATCC 49619	19F	19F	NP	NP	NP	0.25	0.12/2.4 – 1/19	√		√	√
Scotland Pneumococcal Reference Laboratory (Dr. Giles Edwards)	12.1690.X	3	3	NP	NP	180	NP	NP	√			√
	12.1655.E	3	3	NP	NP	180	NP	NP	√			√
	12.1640.H	3	3	NP	NP	180	NP	NP	√			√
	12.1409.G	3	3	NP	NP	180	NP	NP	√			√
	12.1259.V	3	3	NP	NP	180	NP	NP	√			√
	11.6349.S	4	4	NP	NP	205	NP	NP	√			√
	11.5324.N	4	4	NP	NP	246	NP	NP	√			√
	11.4573.Q	4	4	NP	NP	246	NP	NP	√			√
	12.1695.Y	7F	7A/7F	NP	NP	191	NP	NP	√			√
	12.1675.H	7F	7A/7F	NP	NP	191	NP	NP	√			√
	12.1666.R	7F	7A/7F	NP	NP	191	NP	NP	√			√
	12.1426.R	7F	7A/7F	NP	NP	191	NP	NP	√			√
	12.1299.R	7F	7A/7F	NP	NP	191	NP	NP	√			√
	12.1276.C	7F	7A/7F	NP	NP	191	NP	NP	√			√
	11.5715.C	9V	9V	NP	NP	156	NP	NP	√			√
	11.3601.D	9V	9V	NP	NP	162	NP	NP	√			√
	11.3021.E	9V	9V	NP	NP	162	NP	NP	√			√
	11.2143.S	9V	9V	NP	NP	156	NP	NP	√			√
	12.1625.M	19A	19A	NP	NP	2081	NP	NP	√			√
	12.1623.F	19A	19A	NP	NP	416	NP	NP	√			√
	12.1579.M	19A	19A	NP	NP	450	NP	NP	√			√
	10-1688	14	14	NP	NP	NP	NP	NP	√			√
	10-2893	14	14	NP	NP	NP	NP	NP	√			√
	11.4091.H	23F	23F	NP	NP	NP	NP	NP	√			√
	11.3056.E	23F	23F	NP	NP	33	NP	NP	√			√

	11.2827.S	23F	23F	NP	NP	1682	NP	NP	√			√
	11.2737.T	23F	23F	NP	NP	1682	NP	NP	√			√
	11.1373.Z	23F	23F	NP	NP	6959	NP	NP	√			√
	02-4520	5	5	NP	NP	NP	NP	NP	√			√
	04-2077	5	5	NP	NP	1400	NP	NP	√			√
	07-2667	5	5	NP	NP	289	NP	NP	√			√
	10-1351	5	5	NP	NP	4840	NP	NP	√			√
	12.1624.T	6A	6A/6B	NP	NP	396	NP	NP	√			√
	11.6939.E	6A	6A	NP	NP	2467	NP	NP	√			√
	11.5544.Z	6A	6A	NP	NP	1876	NP	NP	√			√
	11.5346.W	6A	6A/6B	NP	NP	327	NP	NP	√			√
	11.3085.J	6A	6A	NP	NP	65	NP	NP	√			√
	10.2671.G	18C	18B/18C	NP	NP	113	NP	NP	√			√
	10.2178.X	18C	18B/18C	NP	NP	113	NP	NP	√			√
	09.2555.S	18C	18B/18C	NP	NP	113	NP	NP	√			√
	09.2223.P	18C	18B/18C	NP	NP	638	NP	NP	√			√
	09.1742.V	18C	18B/18C	NP	NP	2449	NP	NP	√			√
	09.1153.K	18C	18B/18C	NP	NP	1361	NP	NP	√			√
	11.6554.S	19F	19F	NP	NP	NP	NP	NP	√			√
	11.6402.H	19F	1	NP	NP	NP	NP	NP	√			√
GR MICRO	SG07f	7F	7F/7A	NP	NP	NP	NP	NP	√			√
	SG10a	10A	10A	NP	NP	NP	NP	NP	√			√
	SG10f	10F	10F/10C	NP	NP	NP	NP	NP	√			√
	SG11a	11A	11A/11D/18F	NP	NP	NP	NP	NP	√			√
	SG12b	12B	12B	NP	NP	NP	NP	NP	√			√
	SG12f	12F	12B	NP	NP	NP	NP	NP	√			√
	SG15a	15A	15A	NP	NP	NP	NP	NP	√			√
	SG15b	15B	15B	NP	NP	NP	NP	NP	√			√
	SG16f	16F	16F	NP	NP	NP	NP	NP	√			√
	SG17a	17A	17A/34	NP	NP	NP	NP	NP	√			√
	SG18b	18B	18B/18C	NP	NP	NP	NP	NP	√			√
	SG20	20	13/20	NP	NP	NP	NP	NP	√			√
	SG22a	22A	22A/22F	NP	NP	NP	NP	NP	√			√
	SG22f	22F	22A/22F	NP	NP	NP	NP	NP	√			√
	SG23b	23B	23B	NP	NP	NP	NP	NP	√			√

	SG23f	23F	23B	NP	NP	NP	NP	NP	√			√
	SG24b	24B	24B	NP	NP	NP	NP	NP	√			√
	SG24f	24F	24B	NP	NP	NP	NP	NP	√			√
	SG28a	28A	28A	NP	NP	NP	NP	NP	√			√
	SG28f	28F	28A	NP	NP	NP	NP	NP	√			√
	SG29	29	29	NP	NP	NP	NP	NP	√			√
	SG31	31	31	NP	NP	NP	NP	NP	√			√
	SG33b	33B	33B	NP	NP	NP	NP	NP	√			√
	SG33c	33C	35B/35C	NP	NP	NP	NP	NP	√			√
	SG33d	33D	33B	NP	NP	NP	NP	NP	√			√
	SG33f	33F	33A/33F/35A	NP	NP	NP	NP	NP	√			√
	SG34	34	17A/34	NP	NP	NP	NP	NP	√			√
	SG35a	35A	33A/33F/35A	NP	NP	NP	NP	NP	√			√
	SG36	36	7F/21	NP	NP	NP	NP	NP	√			√
Health Protection Agency, Colindale, United Kingdom (Dr. Bruno Pichon, Dr. Jutta Loeffler)	H0 8212 0279	1	1	NP	NP	217	NP	NP	√			√
	H0 8078 4300	1	1	NP	NP	303	NP	NP	√			√
	H0 6274 4730	1	1	NP	NP	618	NP	NP	√			√
	H0 7276 4000	1	1	NP	NP	304	NP	NP	√			√
	H0 8092 8700	1	1	NP	NP	227	NP	NP	√			√
	H0 8282 5540	1	1	NP	NP	228	NP	NP	√			√
	H0 8114 1260	1	1	NP	NP	306	NP	NP	√			√
	H0 8228 5070	1	1	NP	NP	306	NP	NP	√			√
	H0 6196 2040	1	1	NP	NP	2126	NP	NP	√			√
	H0 5122 1380	1	1	NP	NP	3446	NP	NP	√			√
	H0 9260 0327	3	3	NP	NP	180	NP	NP	√			√
	H0 9146 0234	4	4	NP	NP	246	NP	NP	√			√
	H0 8102 0037	4	4	NP	NP	246	NP	NP	√			√
	H0 7406 0041	4	4	NP	NP	246	NP	NP	√			√
	H0 8034 0160	5	5	NP	NP	289	NP	NP	√			√
	H0 8212 0259	6A	6A	NP	NP	65	NP	NP	√			√
	H0 8052 0052	6B	6B	NP	NP	176	NP	NP	√			√
	H0 7156 0309	6B	6B	NP	NP	3481	NP	NP	√			√
	H0 5252 0075	6C	6C/6D	NP	NP	1390	NP	NP	√			√
	H0 8342 0074	8	6B	NP	NP	53	NP	NP	√			√
	H0 9122 0175	8	6B	NP	NP	53	NP	NP	√			√

	H0 7016 0558	9V	9V	NP	NP	156	NP	NP	√			√
	H0 8156 0265	9V	9V	NP	NP	162	NP	NP	√			√
	H0 7174 0058	9N	9N	NP	NP	66	NP	NP	√			√
	H0 7018 0063	9N	9N	NP	NP	66	NP	NP	√			√
	H0 9080 0063	9N	9N	NP	NP	66	NP	NP	√			√
	H0 8208 0041	14	14	NP	NP	9	NP	NP	√			√
	H0 8396 0107	14	14	NP	NP	124	NP	NP	√			√
	H0 8084 0056	14	14	NP	NP	124	NP	NP	√			√
	H0 7442 0047	14	14	NP	NP	124	NP	NP	√			√
	H0 8334 0064	17F	17F/34	NP	NP	964	NP	NP	√			√
	H0 9084 0082	17F	17F/34	NP	NP	392	NP	NP	√			√
	H0 6092 0119	17F	17F/34	NP	NP	392	NP	NP	√			√
	H0 9186 0354	19A	19A	NP	NP	199	NP	NP	√			√
	H0 7336 0087	19A	19A	NP	NP	276	NP	NP	√			√
	H0 8242 0108	19F	19F	NP	NP	162	NP	NP	√			√
	H0 8112 0101	19F	19F	NP	NP	162	NP	NP	√			√
	H0 8432 0293	27	27	NP	NP	1475	NP	NP	√			√

[a] NP: Not Performed

[b] Tanzanian invasive strains: Co-trimoxazole susceptibility determined by Kirby-Bauer disc diffusion test with the following CLSI guidelines (Clinical and Laboratory Standards Institute 2008): susceptible, ≥16 mm, intermediate: 11-15 mm, resistant ≤ 10 mm.

Appendix Table 2. Strains included in Multiple Colonization Study

Child # (Sample Time)	Serotype	Strain ID	Locus allele number							Allelic Profile (ST)	Minimum Inhibitory Concentration (µg/mL)	
			aroE	ddl	gdh	gki	recP	spi	xpt		**Pen**	**Sxt**
1 (Jan/Feb)	NT, no cps PCR	1NP10S	51	14	8	4	10	9	170	4156	0.032	0.19
	NT, no cps PCR	1NP7	51	14	8	4	10	9	170	4156	0.032	0.19
	NT, no cps PCR	1NP4	51	14	8	4	10	9	170	4156	0.032	0.19
	NT, no cps PCR	1NP5	51	14	8	4	10	9	170	4156	0.032	0.19
	NT, no cps PCR	1NP9	51	14	8	4	10	9	170	4156	0.032	0.125
	6B	1NP10B	51	6	66	230	1	6	1	4429	0.19	8
11 (Jan/Feb)	6B	11NP8 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11OP5 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11OP9 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP5 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.25	6
	6B	11OP6 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11OP10 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	8
	6B	11NP2 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP3 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP4 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP6 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP7 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP9 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP10 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11NP1 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
	6B	11OP3 (Jan/Feb)	51	8	66	230	1	6	1	4432	0.19	6
16 (April)	19F	16OP4	12	57	8	9	3	3	20	347	0.19	6
	19F	16NP1	12	57	8	9	3	3	20	347	0.19	6
	19F	16NP4	12	57	8	9	3	3	20	347	0.19	6
	19F	16OP5	12	57	8	9	3	3	20	347	0.19	6
	19F	16OP2	12	57	8	9	3	3	20	347	0.19	6
	19F	16NP3	12	57	8	9	3	3	20	347	0.19	6
	34	16NP5	141	17	5	62	5	9	88	4158	0.032	4
	34	16NP8	141	17	5	62	5	9	88	4158	0.032	4
	6C	16NP9	141	17	5	62	5	9	88	4158	0.032	4
	6C	16NP7	141	17	5	62	5	9	88	4158	0.032	4

	35A	16NP10	1	9	5	36	3	6	1	840	0.25	4
18 (Jan/Feb)	10A	18NP1	16	5	9	75	3	9	115	852	0.125	4
	10A	18NP5	16	5	9	75	3	9	115	852	0.125	4
	10A	18NP9	16	5	9	75	3	9	115	852	0.125	4
	10A	18NP10	16	5	9	75	3	9	115	852	0.125	4
22 (Jan/Feb)	17F	22NP4	10	75	16	54	1	15	1	4160	0.016	0.125
	17F	22NP7	10	75	16	54	1	15	1	4160	0.016	0.125
	17F	22NP9	10	75	16	54	1	15	1	4160	0.016	0.125
	17F	22NP10	10	75	16	54	1	15	1	4160	0.016	0.125
	17F	22NP2	10	75	16	54	1	15	1	4160	0.016	0.125
35 (Jan/Feb)	10A	35NP4	16	5	9	75	3	9	115	852	0.38	6
	19A	35NP10	7	64	11	74	1	6	112	4162	0.125	2
	10A	35NP9	16	5	9	75	3	9	115	852	0.38	6
	6B	35NP1	7	14	9	1	63	9	1	4157	0.125	6
	21	35NP2	8	15	10	2	16	1	1	1145	0.064	3
	6B	35NP6	51	108	66	1	1	6	1	854	0.125	6
	19F	35NP8	12	57	8	9	3	6	20	6170	0.19	2
45 (Jan/Feb)	21	45OP1	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP2	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP3	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP4	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP5	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP6	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP7	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP8	8	15	10	2	16	1	1	1145	0.094	4
	21	45OP9	8	15	10	2	16	1	1	1145	0.094	4
69 (Jan/Feb)	6B	69OP8	51	6	66	230	1	6	1	4429	0.19	8
	6B	69OP9	51	6	66	230	1	6	1	4429	0.19	8
	6B	69OP4	51	6	66	230	1	6	1	4429	0.19	8
	6B	69OP3	51	6	66	230	1	6	1	4429	0.19	8
	6B	69OP7	51	6	66	230	1	6	1	4429	0.19	8
	6B	69OP2	51	6	66	230	1	6	1	4429	0.19	8
	6B	69OP10	51	6	66	230	1	6	1	4429	0.19	6

3 (March)	6B	3NP9 (Mar)	7	14	9	1	10	9	1	4368	0.032	0.125
	6B	3NP6 (Mar)	7	14	9	1	10	9	1	4368	0.032	0.19
	6B	3NP7 (Mar)	7	14	9	1	10	9	1	4368	0.047	0.125
10 (March)	13	10NP1	1	18	11	74	1	6	177	4370	0.125	6
	13	10NP4	1	18	11	74	1	6	177	4370	0.125	6
	13	10NP5	1	18	11	74	1	6	177	4370	0.125	6
	13	10NP7	1	18	11	74	1	6	177	4370	0.125	6
	13	10NP8	1	18	11	74	1	6	177	4370	0.125	6
	13	10NP10	1	18	11	74	1	6	177	4370	0.125	6
11 (June)	6B	11OP1	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP2	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP3	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP4	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP5	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP6	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP8	51	8	66	230	1	6	1	4432	0.19	4
	6B	11OP9	51	8	66	230	1	6	1	4432	0.19	4
3 (September)	1	3OP1	10	9	18	4	1	7	19	217	S	S
	1	3OP2	10	9	18	4	1	7	19	217	S	S
	1	3OP3	10	9	18	4	1	7	19	217	S	S
	1	3OP6	10	9	18	4	1	7	19	217	S	S
	6B	3OP7	1	15	5	54	38	15	288	4373	0.25	6
	6B	3OP8	1	15	5	54	38	15	288	4373	0.25	6
	6B	3OP9	1	15	5	54	38	15	288	4373	0.25	6
	1	3OP10	10	9	18	4	1	7	19	217	S	S
24 (October)	19A	24OP1	7	64	11	4	1	6	112	4162	0.25	6
	19A	24OP3	7	64	11	4	1	6	112	4162	0.25	6
	19A	24OP5	7	64	11	4	1	6	112	4162	0.25	4
29 (November)	35A	29OP8	1	9	5	36	3	6	1	840	0.25	4
	11A	29OP10	2	1	5	36	3	6	339	5752	0.38	6
3 (July)	1	3OP1 (Jul)	10	9	18	4	1	7	19	217	S	S
	1	3OP2 (Jul)	10	9	18	4	1	7	19	217	S	S
	1	3OP4 (Jul)	10	9	18	4	1	7	19	217	S	S
	1	3OP5 (Jul)	10	9	18	4	1	7	19	217	S	S

	1	3OP6 (Jul)	10	9	18	4	1	7	19	217	S	S
10 (Jan/Feb)	19A	10NP2	7	64	11	4	1	6	112	4162	0.125	6
	19A	10NP3	7	64	11	4	1	6	112	4162	0.125	6
	19A	10NP5	7	64	11	4	1	6	112	4162	0.125	6
	19A	10NP6	7	64	11	4	1	6	112	4162	0.125	6
	13	10OP6	1	18	11	74	1	6	177	4370	0.125	3
	19A	10NP7	7	64	11	4	1	6	112	4162	0.125	6
	19A	10NP9	7	64	11	4	1	6	112	4162	0.125	6
10 (April)	18C	10OP1	10	145	11	34	16	15	1	1233	S	S
	18C	10OP3	10	145	11	34	16	15	1	1233	S	S
	18C	10OP5	10	145	11	34	16	15	1	1233	S	S
	18C	10OP6	10	145	11	34	16	15	1	1233	S	S
	18C	10OP7	10	145	11	34	16	15	1	1233	S	S
	6B	10OP9	7	14	9	1	10	9	1	4368	S	S
	18C	10OP10	10	145	11	34	16	15	1	1233	S	S
11 (May)	4	11NP1	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP2	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP3	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP4	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP5	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP6	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP7	ND	ND	ND	ND	ND	ND	ND	ND	S	S
	4	11NP8	ND	ND	ND	ND	ND	ND	ND	ND	S	S

Appendix Table A3. SLVs of STs detected in works included in Chapter 4 and their countries of origins as indicated on MLST database

ST	Countries with ST [a, b]	Related STs	Countries with ST [a, b]
4156	Tanzania (N)	None	-
4429	Tanzania (N)	854	Kenya
		4432	Tanzania (N)
4432	Tanzania (N)	854	Kenya
		4429	Tanzania (N)
347	Norway	556	Sweden
		2715	Kenya
		5028	Mozambique
		5029	Kenya
		5030	Kenya
		5344	Kenya
		5360	Kenya
		5754	Kenya
		5759	Kenya
		5762	Kenya
		5766	Kenya
		5769	Kenya
		5771	Kenya
		5931	Kenya
		6170	Tanzania (N)
		6088	Kenya
		6095	Kenya
		6111	Kenya
		6112	Kenya
		6277	South Africa
4158	Tanzania (N)	None	-
840	Kenya	843	Kenya
		2052	Uganda
		5336	Kenya
		5364	Kenya
		5377	Kenya
		5394	Kenya
		5399	Kenya
		5722	The Gambia
		5723	Kenya
		5724	Kenya
		5772	Kenya
		5902	Kenya
		6315	South Africa
852	Kenya	5304	Kenya
		5329	Kenya
		5876	Kenya
		6098	Kenya
4160	Tanzania (N)	216	Spain
		1146	Kenya
		1794	US
			Nigeria
		3750	Egypt

		5008	South Africa
		5009	India
		5338	Kenya
		5952	Kenya
		6106	Kenya
4162	Tanzania (N)	847	Kenya
			The Gambia
			Spain
		2124	UK
			The Gambia
			Spain
		3326	The Gambia
		5270	Kenya
		5729	The Gambia
		5730	The Gambia
		5820	Malaysia
		5905	Kenya
		5906	Kenya
4157	Tanzania (N)	2265	UK
			Spain
		4368	Tanzania (N)
		5949	Kenya
		5920	Kenya
		6041	Nepal
1145	Kenya	193	International[c]
		375	Norway
		474	UK
			Italy
			Switzerland
		889	Portugal
		1358	Portugal
			Spain
		1888	UK
		2357	Portugal
		2578	Portugal
		2848	Czech
			Republic
		3065	US
		3066	US
		3124	Turkey
		3689	Netherlands
		4021	UK
		4321	UK
		4991	Egypt
4368	Tanzania (N)	2265	UK
		5949	Kenya
		5883	Kenya
		5920	Kenya
4370	Tanzania (N)	701	Kenya
		1144	Kenya
		2069	Uganda
		5340	Kenya
		5922	Kenya

		6109	Kenya
		6110	Kenya
217	International[c]	302	Sweden
		303	US
			Ghana
			Niger
		612	South Africa
			Israel
			Ghana
		613	Kenya
		614	Kenya
		1316	Ghana
		1322	Ghana
		1323	Ghana
		1325	Ghana
		1327	Ghana
		1328	Ghana
		1331	Ghana
			The Gambia
		2019	Egypt
		2034	Egypt
		2206	Niger
		2494	South Africa
		2565	UK
		2772	Syria
		2830	Burkina Faso
		2839	Niger
		2960	Burkina Faso
		3079	Australia
		3081	Oman
		3570	The Gambia
		3575	The Gambia
		3577	The Gambia
		4755	Ghana
		4756	Ghana
		5002	India
		5011-17	South Africa
		5189	Thailand
		5632	Mozambique
		5633	South Africa
		5678	South Africa
		5802	Israel
		5887	Kenya
		5916	Kenya
		5919	Kenya
		6056	Kenya
		6137	Kenya
		6254	Singapore
4373	Tanzania (N)	5248	Kenya
5752	Kenya	843	Kenya
		4984	Mozambique
		5356	Kenya
		5768	Kenya

		5940	Kenya
		5933	Kenya
		5884	Kenya
6170	Tanzania (N)	347	Norway
854	Kenya	5770	Kenya
		5774	Kenya

[a] Countries of origin based on MLST database
[b] N: novel ST

Appendix Table A4. GenBank sequences employed for *cpsB* multiple sequence alignment for works in Chapter 6

Serotype Strain	GenBank Accession Number
1 INV104	FQ312030
1 P1031	NC012467
1 Sp03 2672	FQ312039
1 Sp03 3038	FQ312042
11A AP200	CP002121
14 CGSp14	NC010582
14 INV200	FQ312029
14 JJA	NC012466
19F Taiwan19F-14	NC012469
18C WCH94	AF316642
19A Hungary19A-6	NC010380
19A KEE19A-Ia	JF911512
19A KEE19A-Ib	JF911513
19A KEE19A-Ic	JF911516
19A KEE19A-Id	JF911517
19A KEE19A-Ie	JF911518
19A KEE19A-If	JF911519
19A KEE19A-Ig	JF911521
19A KEE19A-II	JF911514
19A KEE19A-IIIa	JF911511
19A KEE19A-IIIb	JF911520
19F KEE19F-IIIa	JF911522
19A TCH8431	CP0019931
19F G54	NC011072
19F KEE19F-Ia	JF911525
19F KEE19F-Ib	JF911526
19F KEE19F-Ic	JF911527
19F KEE19F-Ie	JF911531
19F KEE19F-IIb	JF911529
19F KEE19F-IIc	JF911530
19F KEE19F-IIIb	JF911523
19F KEE19F-IVa	JF911528
19F NCTC11906	AF030367
19F PO329	AF030371
19F SP496	AF030368
19F SPGA31	AF030370
19F SPVA92	AF030369
19F SPVA96	AF030372
2 D39	CP000410
23F ATCC700669	FM211187
23F UK-577	AF030374
23F SP-264	AF030373
3 OXC141	FQ312027
3 Sp99 4038	FQ312041
3 Sp99 4039	FQ312044
3 SPN034156	FQ312045
3 SPN034183	FQ312043
37 7077-39	AJ131984
4 TIGR4	AE005672
4 WCH35	WCH35
5	NC012468
5 70585	CP000918
6A KEE6A-Ia	JF911487
6A KEE6A-IIa	JF911488
6A KEE6A-IIIa	JF911489
6A KEE6A-VI	JF911490

6A-KEE6A-Ib	JF911496
6A-KEE6A-IIb	JF911488
6A-KEE6A-IIIb	JF911489
6A-KEE6A-IVa	JF911492
6A-KEE6A-IVb	JF911494
6A-KEE6A-Va	JF911491
6A-KEE6A-Vb	JF911495
6B 670-6B	NC014498
6B PN94/361	AF246897
6B WCH18	AF316640
6B-KEE6B-1c	JF911501
6B-KEE6B-Ib	JF911500
6B-KEE6B-Id	JF911502
6B-KEE6B-Ie	JF911503
6B-KEE6B-If	JF911506
6B-KEE6B-Ig	JF911508
6B-KEE6B-IIa	JF911499
6B-KEE6B-IIb	JF911505
6B-KEE6B-IIIa	JF911504
6B-KEE6B-IIIb	JF911507
6B-KEE6B-Ia	JF911498
6C CHPA388	EF538714
6C-KEE6C-Ia	JF911509
6C-KEE6C-Ib	JF911510
6C-KEE6C-Ic	JF911515
6D MNZ21	HM171374
6D MNZ920	HM448897
8 WCH56	AF316641
1 519/43	CR926497
10A 10061/38	CR931649
10B 423/82	CR931650
10C Gro Norge	CR931651
10F 34355	CR931652
11A 1813/39	CR931653
11B 8087/40	CR931654
11C Eddy nr. 53	CR931655
11D 70/86	CR931656
11F 34356	CR931657
12A 559/66	CR931658
12B Gambia 1/81	CR931659
12F 6312	CR931660
13 34357	CR931661
14 34359	CR931662
15A 389/39	CR931663
15B 7904/39	CR931664
15C 533/62	CR931665
15F 688/63	CR931666
16A R105	CR931667
16F nr.34361	CR931668
17A nr.4704	CR931669
17F Rose	CR931670
18A 8609/43	CR931671
18B 1033/41	CR931672
18C 4593/40	CR931673
18F Gethens	CR931674
19A 141/68	CR931675
19B nr.4594	CR931676
19C 7588/39	CR931677
19F 485/61	CR931678
2 pN2L	CR931633
20 34365	CR931679

21 546/62	CR931680
22A 3405/39	CR931681
22F 1772/40	CR931682
23A 1196/45	CR931683
23B 1031/41	CR931684
23F Dr. Melchior	CR931685
24A 2748/40	CR931686
24B 2236/42	CR931687
24F L	CR931688
25A tp 25/38	CR931689
25F 601/62	CR931690
27 nr.34371	CR931691
28A 1982/45	CR931692
28F 34372	CR931693
29 nr.34373	CR931694
3 542/62	CR931634
31 nr.34374	CR931695
32A 2813/41	CR931696
32F nr.34375	CR931697
33A Biehl	CR931698
33B E294	CR931699
33C 7098/41	CR931700
33D CSF/79	CR931701
33F 3084/37	CR931702
34 676/74	CR931703
35A 1936/39	CR931704
35B 4356/39	CR931705
35C 7765/43	CR931706
35F 361/39	CR931707
36 1095/39	CR931708
37 264/73	CR931709
38 9687/39	CR931710
39 203/40	CR931711
4 600/62	CR931635
40 Colemore	CR931712
41A 6803	CR931713
41F 8211/40	CR931714
42 198/71	CR931715
43 2427/48	CR931716
44 Hammer	CR931717
45 Eddy nr.72	CR931718
46 Eddy nr.73	CR931719
47A L351	CR931720
47F Eddy nr.52	CR931721
48 656/63	CR931722
5 Ambrose	CR931637
6A 34351	CR931638
6B 2616/39	CR931639
7A 2040/37	CR931640
7B Johnson	CR931641
7C Sutcliff	CR931642
7F 554/62	CR931643
8 573/62	CR931644
9A Wilder	CR931645
9L T9233/128/68	CR931646
9N 533/62	CR931647
9V 980/68	CR931648

TPC Limited

P.O. BOX 93
MOSHI
KILMANJARO
TANZANIA
TEL: +255 27 2754389
FAX: +255 27 2754391
E-MAIL:tpc@tpc.co.tz

Ms. Ndekya M. Oriyo
Ms. Emily Bongard
KCMC
P.O. Box 3010
MOSHI

Your ref.:
Our ref.: L08E02RB

Date: 2 May 2008

Dear Madam,

Re: PERMISSION TO DO RESEARCH

The TPC Ethical Committee is happy to inform you that on its sitting 30/4/08, after scrutinizing your application to do Research on Sereotype Distribution of Carriage & Invasive Pneumococcal Isolates from children at TPC, has given its consent with the following requests.

(a) In capacity building we request for training of one of our laboratory employee.

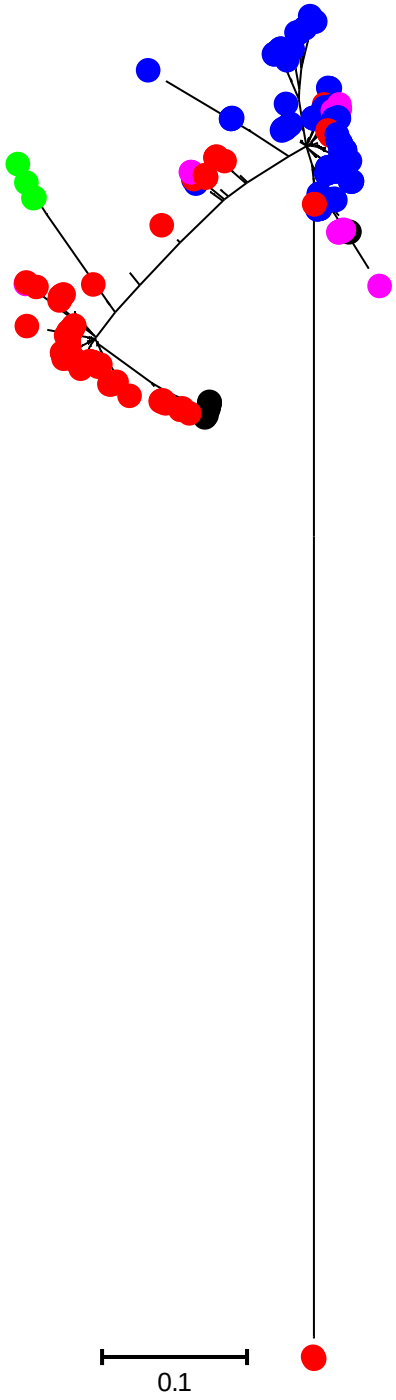
(b) Give feedback after the research completion.

The Committee congratulates you with best wishes.

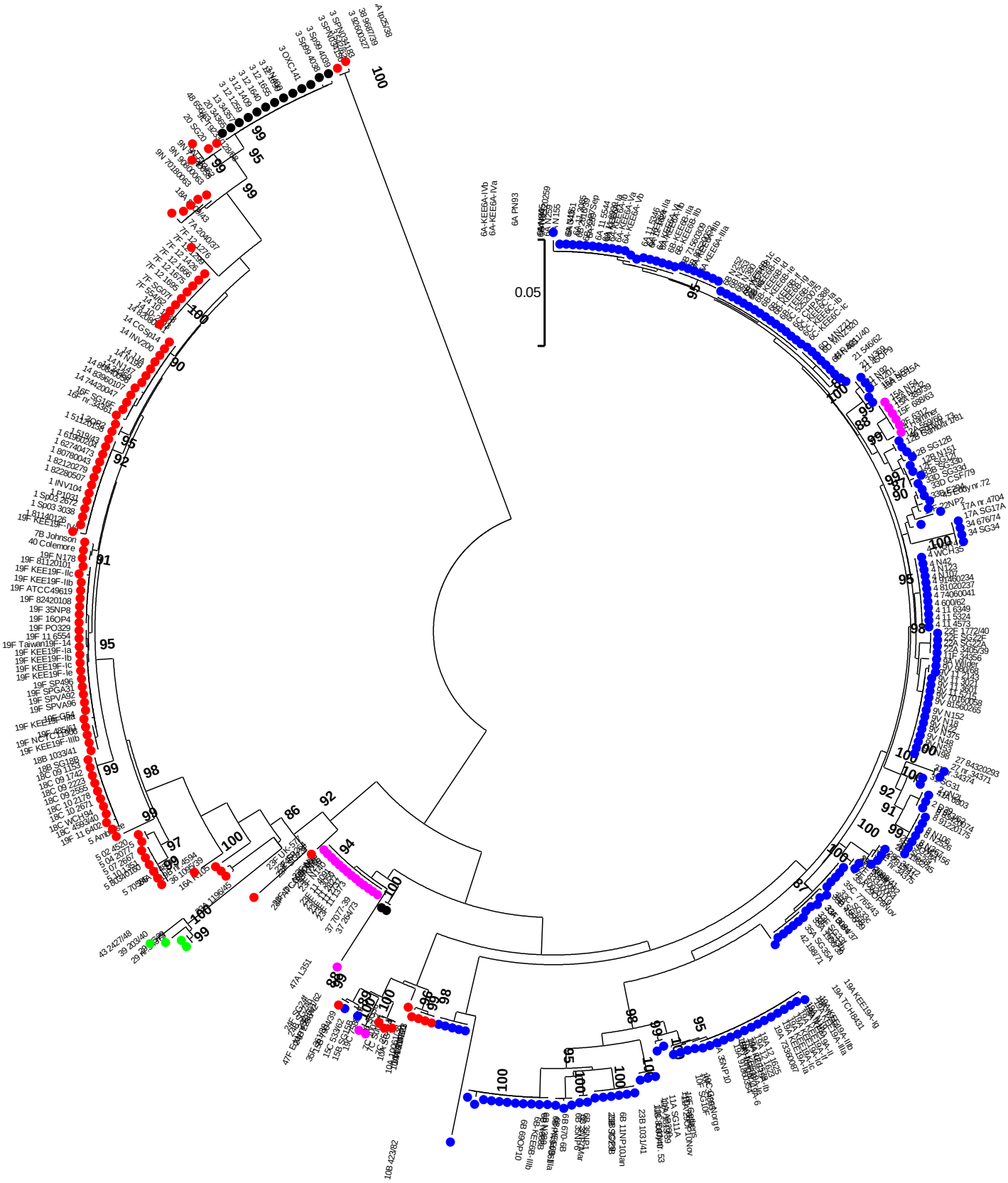
Secretary
TPC ETHICAL COMMITTEE

Chairperson
TPC ETHICAL COMMITTEE

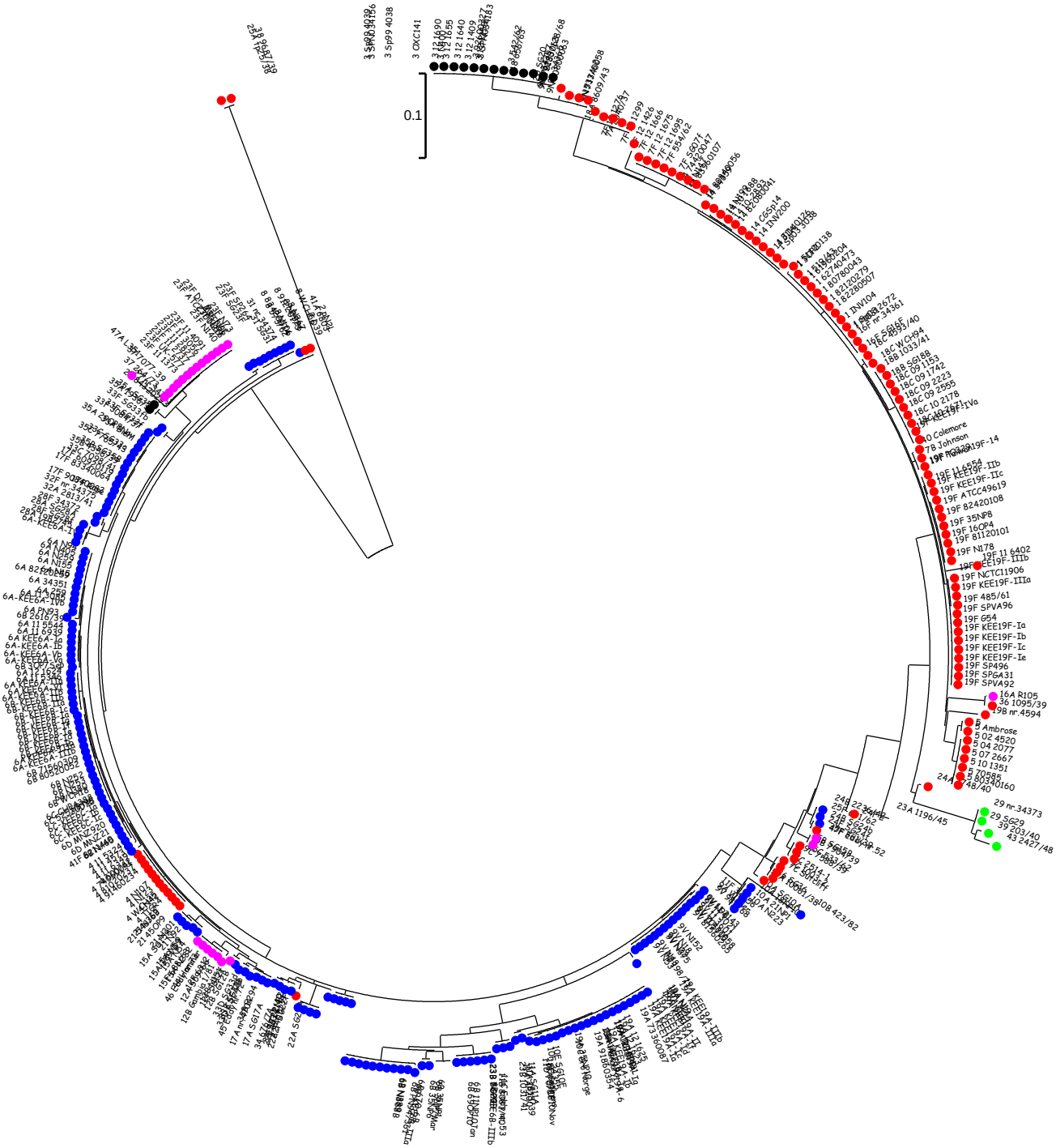
Appendix Fig. A1. Ethical approval document granted by TPC



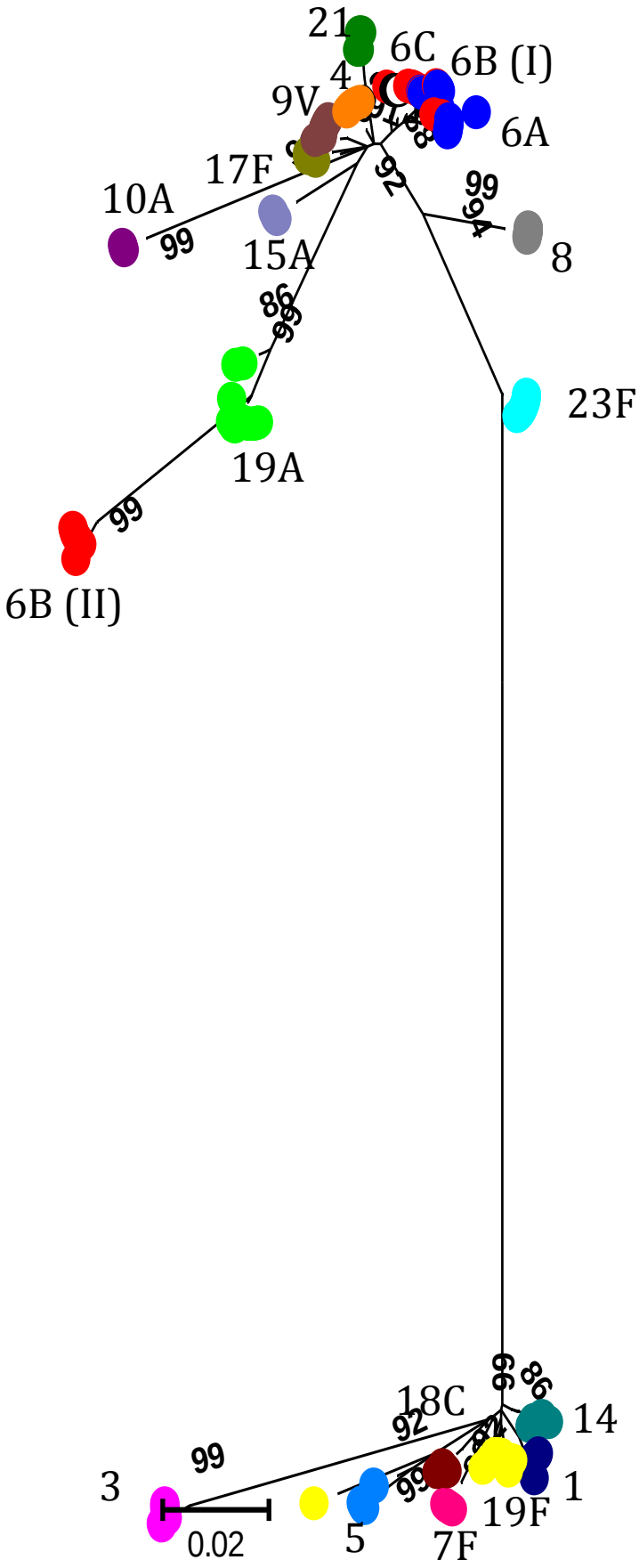
Appendix Fig. A2. Maximum-likelihood phylogenetic representation of pneumococcal *cpsB* sequences. Red and blue colouring was based on study of Varvio *et al.* (Varvio, 2009), where blue represents serotypes commonly associated with carriage, and red represents those predominantly found in invasive diseases. Serotypes which, according to Varvio *et al.*, had blue-red mosaic within *cpsB* (serotypes 15A, 15F, 23F, 35F, 47A, 47F) are indicated in purple. Serotypes 29, 39, and 43 contained mosaic sequences resembling red sequences and other streptococcal sequences, and are indicated in green. Serotypes 3 and 37 do not require *cpsB* for capsule production, and are indicated in black. Phylogenetic representation performed on MEGA 5.05 with bootstrapping of 100 repetitions. Distances represented as the number of nucleotide substitutions per site. A total of 351 sequences were included in analysis.



Appendix Fig. A3. Neighbour-joining phylogenetic (radial) representation of pneumococcal *cpsB* sequences. Red and blue colouring was based on study of Varvio *et al.* (Varvio *et al.* 2009), where blue represents serotypes commonly associated with carriage, and red represents those predominantly found in invasive diseases. Serotypes which, according to Varvio *et al.*, had blue-red mosaic within *cpsB* (serotypes 15A, 15F, 23F, 35F, 47A, 47F) are indicated in purple. Serotypes 29, 39, and 43 contained mosaic sequences resembling red sequences and other streptococcal sequences, and are indicated in green. Serotypes 3 and 37 do not require *cpsB* for capsule production, and are indicated in black. Phylogenetic representation performed on MEGA 5.05 with bootstrapping of 1,000 repetitions, with bootstrapping values over 85% indicated. Distances represented as the number of nucleotide substitutions per site. A total of 351 sequences were included in analysis.



Appendix Fig. A4. Maximum-likelihood phylogenetic (radial) representation of pneumococcl *cpsB* sequences. Red and blue colouring was based on study of Varvio *et al.* (Varvio *et al.* 2009), where blue represents serotypes commonly associated with carriage, and red represents those predominantly found in invasive diseases. Serotypes which,according to Varvio *et al.*, had blue-red mosaic within *cpsB* (serotypes 15A, 15F, 23F, 35F, 47A, 47F) are indicated in purple. Serotypes 29, 39, and 43 contained mosaic sequences resembling red sequences and other streptococcal sequences, and are indicated in green. Serotypes 3 and 37 do not require *cpsB* for capsule production, and are indicated in black. Phylogenetic representation performed on MEGA 5.05 with bootstrapping of 100 repetitions. Distances represented as the number of nucleotide substitutions per site. A total of 351 sequences were included in analysis.



Appendix Fig. A5. Maximum-likelihood phylogenetic representation of serotypes with more than five strains each. Serotypes are indicated beside clusters of taxa. Bootstrap with replacements was performed with 100 repetitions, and bootstrap values above 85% ware indicated. Ruler indicates average number of nucleotide differences per site. Phylogenetic analysis performed on MEGA 5.05.

Infection, Genetics and Evolution 11 (2011) 1989–1995

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

The adaptive potential during nasopharyngeal colonisation of *Streptococcus pneumoniae*

Marcus H.Y. Leung [a], Ndekya M. Oriyo [a,b], Stephen H. Gillespie [a,c,*], Bambos M. Charalambous [a]

[a] Centre for Clinical Microbiology, University College London Medical School, Royal Free Campus, Rowland Hill Street, London NW3 2PF, United Kingdom
[b] Kilimanjaro Christian Medical College, P.O. Box 2240, Moshi, Tanzania
[c] School of Medicine, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 9TF, United Kingdom

ARTICLE INFO

Article history:
Received 22 July 2011
Received in revised form 31 August 2011
Accepted 3 September 2011
Available online 10 September 2011

Keywords:
Streptococcus pneumoniae
Multiple colonisation
Multilocus sequence typing
Adaptation
Genetic diversity
Nasopharyngeal carriage

ABSTRACT

Adaptation to host defences and antimicrobials is critical for *Streptococcus pneumoniae* (the pneumococcus) during colonisation of the nasopharynx – its only ecological habitat. The pneumococcus is highly transformable with the genome between different strains varying widely in both gene content and gene sequence. Thus, mixed strains colonising together will expand the genetic reservoir – "supragenome" for this highly transformable microorganism, increasing its adaptive potential.

The extent of the phenotypic and genotypic diversity of strains co-colonising in the nasopharynx was determined. In contrast to most carriage studies, which characterise single colonies, a systematic analysis of up to 20 colonies per colonisation was undertaken in Tanzanian children for 12 months. The serotype was determined by conventional serology and confirmed by DNA-based methods. The antibiotype for penicillin and co-trimoxazole was determined from the minimum inhibitory concentration determined by *E*-test. As representative of the genotype of strains the sequence types (STs) was determined by multilocus sequence typing (MLST).

Of 61 colonisation events studied, seven (11.5%) had strains expressing multiple serotypes, with a maximum of five serotypes detected. Four colonisation events also had co-colonisation of penicillin and/or co-trimoxazole susceptible and non-susceptible pneumococci. Sequence typing revealed that 58% were unique to our cohort. Simultaneous colonisation of up to six STs with two expressing serotype 6B was seen. Re-isolation of either the same or different strains of serotype 6B was seen. Genetically related single-locus and double-locus variants were identified in our cohort that differed by multiple nucleotides.

Multiple colony characterisation revealed phenotypic and genetic evidence of microevolution and a greater diversity of pneumococcal strains colonising together than previously observed, thus increasing the potential to adapt in response to the host environment during colonisation.

1. Introduction

Streptococcus pneumoniae (the pneumococcus) is a major cause of bacterial otitis media, sinusitis, pneumonia, septicaemia, and meningitis with annual mortality in children approaching a million (World Health Organization, 2007). Despite its pathogenicity, the pneumococcus predominantly colonises the nasopharynx, with high prevalence in asymptomatic children due to their immature immune system (Bogaert et al., 2004). Pneumococcal colonisation prevalence varies by geographical regions, and studies in Africa have revealed pneumococcal colonisation rates of over 80% in children under the age of five (Huebner et al., 1998; Adegbola et al., 2001).

The pneumococcus is a highly transformable organism, which allows the uptake of DNA from the extracellular environment. Whole-genome analysis reveals extensive genetic variations between colonising pneumococcal strains with as little as 46% of orthologous gene clusters conserved (Shen et al., 2006; Hiller et al., 2010). From these data and previous studies on other naturally-transformable organisms a model of the increased size of the pneumococcal "supragenome" that may be available for adaptation when multiple strains are colonising the nasopharynx was developed (Hiller et al., 2007; Tettelin et al., 2005; Hogg et al., 2007). Co-colonisation of different strains thus allows acquisition of novel genes from an enlarged gene pool, potentially contributing to adaptation and evolution. Colonisation of multiple strains differing in phenotypes, such as the polysaccharide capsule, which has more than 90 different serotypes (Bentley et al., 2006), allows continued survival by evasion of serotype-specific immune responses as well as adaptation by genomic alteration. For all serotypes, the

* Corresponding author. Present address: School of Medicine, University of St. Andrews, North Haugh, St. Andrews, Fife KY16 7TF, United Kingdom.
E-mail address: shg3@st-andrews.ac.uk (S.H. Gillespie).

1567-1348/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2011.09.002

capsulation locus is flanked by conserved genes *dexB* and *aliA* that enable serotype switching by homologous recombination of the entire capsulation cassette (Coffey et al., 1998; Brueggemann et al., 2007). In addition, horizontal gene transfers (HGTs) of antibiotic-resistance determinants, such as genes encoding the penicillin-binding proteins (Laible and Hakenbeck, 1991; Dowson et al., 1993; Hakenbeck et al., 1998), are crucial in the emergence of drug resistance.

Despite the biological importance of colonisation by multiple pneumococcal strains, there is a paucity of data on multiple pneumococcal colonisations. It is known that there is competition between different serotypes using a mouse colonisation model (Lipsitch et al., 2000) as well as studies in asymptomatic children in day-care centres (Auranen et al., 2010). However, these studies do not describe the biological significance of co-colonisation of multiple serotypes and gentoypes. Most carriage studies characterise only a single colony within a colonisation event. While this provides an accurate index of the prevalent strains circulating within a community, single colony characterisation overlooks the detailed ecology within the nasopharynx (Charalambous et al., 2008). In the few studies where multiple colonies have been characterised (Hansman et al., 1985; Gratten et al., 1989; Huebner et al., 2000; Meats et al., 2003; Bronsdon et al., 2004; O'Brien et al., 2007; Hare et al., 2008; Hill et al., 2008; Kaltoft et al., 2008; Auranen et al., 2010), these studies showed large variation of co-colonisation rates (from 1.2% to 33.3%), with a maximum of three serotypes within a colonisation event. In addition, these methods are highly variable with limited data on the phenotype and genotype reported (Meats et al., 2003; Hill et al., 2008; Vestrheim et al., 2008). Colonising strains with identical phenotypes can be genetically different (Dagerhamn et al., 2008; Donkor et al., 2011), and studying multiple colonisation diversity by phenotypes alone will certainly underestimate the true heterogeneity of colonising strains. Surprisingly, only two pneumococcal carriage studies to our knowledge have characterised the serotype, antibiotype, and genotype (Sá-Leão et al., 2002; Vestrheim et al., 2008).

Thus, to unravel the detailed ecology of the pneumococcus a systematic analysis of up to 20 colonies from each colonisation event in Tanzanian children living on a semi-isolated sugar plantation was undertaken over a 12-month period. The serotype (as determined by serological and confirmed by DNA-based methods) and antibiotic susceptibilities to penicillin and co-trimoxazole were characterised. Multilocus sequence typing (MLST) was performed on all the multiple serotype colonisations and a similar number of randomly selected single serotype colonisation events to determine both phenotypic and genetic diversity of colonising strains.

2. Materials and methods

2.1. Ethical approval and demographic characteristics

Ethical approval was granted by the Tanganyika Planting Company (TPC) Limited Ethical Committee. The study site is a remote community located within a sugar plantation, TPC, located in north-eastern Tanzania 25 km from Moshi. The 2300 employees and their families who live on the plantation are served by the 100-bed TPC Hospital with free point-of-care treatment. The plantation is divided into camps and living accommodation is allocated according to occupation and income. Most children in the study are from families with typical socioeconomic status in Tanzania, with average family size of five members. Their homes have on average two bedrooms with shared bathroom and toilet facilities. Firewood is used for cooking in greater than 90% of the households. Every camp has a small drug dispensary.

2.2. Recruitment criteria

Healthy children aged 1–5 years of age attending the Mother and Child Health Centre at TPC Hospital for routine examination and the expanded programme of immunization (EPI) vaccination (pneumococcal vaccine not included) were recruited following informed consent from the parent or guardian. The recruitment criteria for the children were that they resided on the plantation, were healthy with no antibiotics use over the 3 months prior to recruitment. A standard pro-forma was completed by the investigator or clinician giving details of the child's gender, age, medical history, and current antibiotic treatment. It included a continuing record of the antibiotics and other medicines taken within a month of each sampling point. Eligible children positive for pneumococcal carriage were swabbed monthly, where possible, for the following 12 months. Recruitment of children was initiated in January and February 2003 and followed up until December 2003.

2.3. Specimen collection and detection of multiple colonisations

Nasopharyngeal and oropharyngeal swabs were taken for the first 6 months of the study. Due to difficulties encountered in repeat sampling, only throat swabs were taken for the remaining 6 months. Calcium-alginate swabs on a flexible aluminium shaft and cotton tipped swabs were used for nose and throat swabs, respectively. The swab was immediately immersed in 1.0 mL of STGG (skim milk-tryptone-glucose-glycerol) medium contained in a 7.0 mL plastic screw-topped bijoux bottle (O'Brien et al., 2001). The shaft of the swab was cut with ethanol-flamed, sterilized scissors approximately an inch above the swab to allow the lid to be replaced immediately. The transfer was performed quickly to keep contaminants to a minimum. Swabs were taken to the Kilimanjaro Christian Medical Centre (KCMC) referral hospital where they were stored at −70 °C. Ten microlitres of the archived sample was plated onto a Columbia blood agar plate made selective by supplementing with 5 μg/mL gentamicin. Up to 20 colonies from the initial screen (10 from nose and 10 from throat swabs) were purity plated on blood agar. Where morphologically different colonies were seen these were selected, and the final number of colonies sub-cultured for purity depended on the number observed on the initial plate. Pneumococci were identified by α-haemolysis, colony morphology, optochin sensitivity, and bile solubility. Pneumococcal isolates were frozen in STGG medium and shipped on dry ice to our London laboratory for full characterisation.

2.4. Serotyping

Pooled antisera using the chessboard method (Sørensen, 1993) following the manufacturer's protocol (Statens Serum Institut, Copenhagen, Denmark) were used. When multiple serotypes were observed within a colonisation, serotyping was confirmed by sequencing *cpsB* on three independent clones, and comparing to *cpsB* sequences of known serotypes made available by Wellcome Trust (Bentley et al., 2006, Leung, in preparation). Additionally, pneumococci typed as either serotype 13/28 and/or 17A/34 were confirmed by multiplex PCR methods described by Kong et al. (2005) and Pai et al. (2006), respectively. Strains that failed to agglutinate with antisera were defined as acapsulate if no amplicons of the *cpsA* and *cpsB* genes were observed (Pai et al., 2006, Leung, in preparation).

2.5. Antibiotic susceptibility

Susceptibility to penicillin, co-trimoxazole (Sxt), tetracycline, chloramphenicol, erythromycin, and amoxicillin was performed by disc diffusion (Clinical Laboratory and Standards Institute,

2010). *E*-test was performed on non-susceptible strains to penicillin with benzylpenicillin strips and to co-trimoxazole (AB Biodisk, Sölna, Sweden). *S. pneumoniae* ATCC 49619 was used as the control strain. Antibiotypes were defined as different only if they had a minimum inhibitory concentration (MIC) that varied by at least fourfold. MIC determinations were performed on three independent clones of strains within an individual colonisation event with different antibiotypes. Where the MIC was performed in triplicate or where multiple strains within a colonisation event showed the same antibiotype, the median is given.

2.6. Preparation of genomic DNA

Archived pneumococcal isolates were sub-cultured onto Columbia blood agar (Oxoid) and incubated overnight at 37 °C with 5% CO_2. Genomic DNA was extracted by the heat lysis method. A sweep of cells is resuspended in 50 μL of phosphate-buffered saline, pH 7.4 (Gibco) and heated to 95 °C for 5 min. The cell debris is pelleted in a microfuge at 10,000*g* for 5 min. The supernatant containing genomic DNA was diluted 1:10 in PBS (pH 7.2, Gibco) before MLST analysis.

2.7. Multilocus sequence typing (MLST) and eBURST analysis

MLST was performed by nucleotide sequence analysis of loci in seven housekeeping genes as previously described (Enright and Spratt, 1998). Amplicons were purified using the QIAquick PCR Purification Kit (QIAGEN) and sequenced using the Big Dye Sequence Terminator Kit (Applied Biosystems). Samples were run on 3130 Genetic Analyser (Applied Biosystems) and analysed using the Bionumerics software (Applied Maths). The pneumococcal MLST database (http://www.mlst.net) was interrogated with sequence data for each locus and allelic numbers were assigned when homologous alleles were matched. Unassigned alleles were submitted to the curator of the pneumococcal MLST database. From the seven-integer allele profile a sequence type (ST) was assigned. New STs were submitted to the curator of the MLST database. eBURST (http://www.eburst.mlst.net) was used to determine the genetic relatedness of different sequence types (Feil et al., 2004). Genetically related strains are defined as sharing 6/7 (single-locus variant; SLV) or 5/7 (double-locus variant; DLV) MLST loci. SLV and DLV pairs were confirmed by repeating the MLST analysis on three independent clones. The chromatograms were also independently verified.

3. Results

Of 83 children recruited to the study, 21 (25%) were colonised by *S. pneumoniae* and were followed up monthly. The mean age of the 21 children was 3.2 ± 1.6 (Standard Deviation, SD) years. On average 14 ± 4.4 (67%) of the children were swabbed per month. Each child was swabbed on average 7.0 ± 2.6 times and was positive for pneumococci on 3.0 ± 1.2 occasions. A total of 61 pneumococcal colonisation episodes were captured. An average of 6.7 ± 1.7 colonies were sub-cultured and characterised per colonisation event. During follow-up there was exposure to antibiotics in the previous month of sampling in 44% (27/61) of the pneumococcal colonisation events. There was no significant difference in the colonisation rates, or the level of antibiotic non-susceptible strains (data not shown) in children on antibiotics compared with those not on antibiotics.

Seven (12%) pneumococcal colonisation events from six children had multiple serotypes. Four of these also had multiple antibiotypes. Seventeen of the 61 pneumococcal colonisations, including the seven multiple strain colonisations and 10 randomly chosen single serotype colonisations were sequence-typed by MLST.

3.1. Colonisation by multiple serotypes of S. pneumoniae

Thirteen serotypes (1, 4, 6B, 10A, 11A/D, 13, 17F, 18C, 19A, 19F, 21, 34, and 35A) were detected within the 17 colonisation events analysed by sequence typing in addition to phenotype characterisation. A non-serotypable strain was captured in one event. In the seven multiple phenotype colonisation events identified, up to five serotypes (6B, 10A, 19A, 19F, 21) were detected colonising together (Table 1). The most common serotype observed within the multiple phenotype colonisation events studied was 6B (4/7, 57%) followed by 19A, 19F and 35A, each found in two events (2/17, 29%). The proportion of each serotype in mixed colonisations was investigated. A dominant serotype (defined in this study as >80% prevalence within a colonisation) was found in 43% (3/7) colonisation events, with each having two serotypes (Table 1). A dominating serotype was not observed for colonisation events with more than two serotypes, but serotype 35A made up less than 10% of the population characterised in Child 16, who carried three serotypes (Table 1).

3.2. Colonisation by multiple antibiotypes of S. pneumoniae

Multiple antibiotypes with at least a fourfold difference in MIC were detected in four colonisations, all of these also had multiple serotypes (Table 1). All four events had penicillin susceptible and non-susceptible strains colonising together, while two events were also colonised simultaneously by Sxt-susceptible and non-susceptible strains (Table 1). Serotype 6B was associated with multiple antibiotypes, differing in both penicillin and Sxt susceptibilities (Table 1, Child 11 January, Child 3 March). In Child 29, serotypes 35A and 6A were detected with the same antibiotype for both penicillin and Sxt. In children where multiple colonisation events were found on more than one occasion (Child 3, Child 11), strains of the same serotype did not have a different antibiotype (Table 1).

3.3. Colonisation by multiple sequence types (STs) of S. pneumoniae

Sequence typing was determined on 17 colonisation events to delineate genotype diversity. A total of 19 STs were identified, eleven of which (57.9%) had not been described previously. Of the eight STs previously-reported, five (63%) were reported to have been isolated in Africa, and the remaining three (38%) STs, namely ST217, ST347, and ST1233, were isolated in both Europe and Africa. Multiple STs were detected in all seven colonisation events with multiple serotypes (Table 1). Up to six STs, of which two with unrelated STs (that shared only two MLST loci) had the same serotype (6B) were detected within a single colonisation in Child 35 (Table 1). Several STs, namely ST840, ST852, ST1145, and ST4162 were found in more than one child (Table 1). ST217 expressing serotype 1 was found in two colonisation events in Child 3 2 months apart (Table 1). ST4432 was also found in two colonisation events in Child 11 6 months apart (Table 1). Multiple STs with the same serotype were found (Table 2).

3.4. Repeated colonisation of strains with the same serotype

On three occasions repeated colonisation of the same serotype in the same child was observed from 2 to 5 months in duration, but it is difficult to resolve whether this is continued carriage with the level of carriage below the limit of detection, or re- acquisition (Table 1). Another method with greater sensitivity may be able to differentiate between them. In contrast, in Child 3 a different strain of 6B was isolated 5 months later (Table 1).

Table 1
The serotypes, antibiotypes, and sequence types of pneumococcal strains characterised in this study.

Child no.	Colonisation (sampling month)	No. of isolates characterised	Serotype	MIC[a]		Sequence type	No. (%)
				Penicillin	Sxt		
1	January/February	6	Acapsular[b]	0.032	0.19	4155	5 (83)
			6B	0.19	8	4429	1 (17)
3	March	3	6B	0.032	0.125	4368	3 (100)
3	July	5	1	<0.016	<0.002	217	5 (100)
3	September	8	1	<0.016	<0.002	217	5 (63)
			6B	0.25	6	4373	3 (38)
10	January/February	7	13	0.125	3	ND	1 (14)
			19A	0.125	6	ND	6 (86)
10	March	6	13	0.125	6	4370	6 (100)
10	April	7	18B	<0.016	<0.002	ND	6 (86)
			6B	<0.016	<0.002	ND	1 (14)
11	January/February	15	6B	0.19	6	4432	15 (100)
11	May	8	4	<0.016	<0.002	ND	8 (100)
11	June	8	6B	0.19	4	4432	8 (100)
16	April	11	19F	0.19	6	347	6 (55)
			34	0.032	6	4158	4 (36)
			35A	0.25	4	840	1 (9)
18	January/February	3	10A	0.125	4	852	3 (100)
22	January/February	5	17F	0.016	0.125	4160	5 (100)
24	October	3	19A	0.25	6	4162	3 (100)
29	November	2	35A	0.24	4	840	1 (50)
			11A/D	0.38	6	5752	1 (50)
35	January/February	7	6B	0.125	6	854	1 (14)
				0.125	6	4157	1 (14)
			10A	0.38	6	852	2 (29)
			19A	0.125	2	4162	1 (14)
			19F	0.19	2	6170	1 (14)
			21	0.064	3	1145	1 (14)
45	January/February	9	21	0.094	4	1145	9 (100)
69	January/February	7	6B	0.19	8	4429	7 (100)

ND (no data as MLST was not performed).
[a] MIC, minimum inhibitory concentration, as the median of triplicate results.
[b] Strains were nontypeable by pooled sera with absence of *cpsA–cpsB*.

Table 2
Serotypes with multiple sequence types detected within our cohort.

Serotype	*aroE*	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*	ST	Child	Colonisation event
6B	51	66	230	1	6	1	8	4432	11	January/February
	51	66	230	1	6	1	8	4432	11	June
	7	9	1	10	9	1	14	4368	3	March
	1	5	54	38	15	288	15	4373	3	September
	51	66	1	1	6	1	108	854	35	January/February
	7	9	1	63	9	1	14	4157	35	January/February
	51	66	230	1	6	1	6	4429	1	January/February
	51	66	230	1	6	1	6	4429	69	January/February
19F	12	8	9	3	3	20	57	347	16	April
	12	8	9	3	6	20	57	6170	35	January/February

3.5. Genetic relatedness of sequence types

Genetically related strains were defined as either single-locus variants (SLVs) with 6/7 identical loci, or double-locus variants (DLVs) with 5/7 identical loci. All but seven of the STs were unrelated. Three pairs of SLVs were identified in different colonisations from different children (Table 1). ST6170 and ST347 (both serotype 19F and the same antibiotype), differed in the *spi* locus by a single nucleotide (Table 2) ST4432 and 4429 (both serotype 6B and the same antibiotype) differed in the *ddl* locus by four nucleotides (Table 2). ST4157 and ST4368 differed in the *recP* locus by three nucleotides (Tables 1 and 2). ST854 (serotype 19F) is a DLV of ST4432 (serotype 6B) and ST4429 (6B), and differed at the *gki* locus by five nucleotides and the *ddl* locus by 40 nucleotides (Table 2). All three STs have the same antibiotype (Table 1). Allele 108 of the *ddl* locus has interspecies homology to *ddl* of *Streptococcus mitis* (Accession Numbers EU075717.1 and EU075693.1) and *Streptococcus oralis* (AJ387988.1) (Table 3).

4. Discussion

Colonisation by multiple pneumococci provides a greater potential for successful phenotypes to colonise as well as providing the framework for genetic exchange between strains that differ in gene alleles and gene contents. Such a phenomenon enlarges what has been described as the "supragenome" of naturally transformable species (Shen et al., 2005, 2006; Tettelin et al., 2005; Hiller et al., 2007, 2010; Donati et al., 2010). It is important, therefore, to study the degree of multiple colonisations amongst children to understand and define the potential for genetic exchange. To achieve this we characterised a random selection of up to 20 colonies from pneumococcus-positive nasopharyngeal samples. Two critical phenotypes, the serotype and antibiotic susceptibility, as well as the sequence type were determined.

Our data reveals the extensive and so far unrecognized potential for genetic ex-change in the nasopharynx of children in Tanzania. Nineteen strains were detected in 12% of the colonisation

Table 3
Allelic variation between single and double locus variants detected within our cohort.

Serotype	SLV		Variant locus	Variant alleles (A/B)	Nucleotide substitutions (% identity)
	Sequence type A	Sequence type B			
19F	347	6170	*spi*	3/6	1 (99.8)
6B	4432	4429	*ddl*	8/6	4 (99.1)
6B	4157	4368	*recP*	63/10	3 (99.3)
	DLV				
6B	854	4432	*ddl*	108/8	38 (91.4)
			gki	1/230	5 (99.0)
6B	854	4429	*ddl*	108/6	40 (90.9)
			gki	1/230	5 (99.0)

episodes with extensive serotype, antibiotype and ST heterogeneity within a colonisation event that had the potential to utilise the supragenome. This included the presence of antibiotic-susceptible and non-susceptible strains colonising together. The co-colonisation of two unrelated strains expressing the same serotype 6B is reported for the first time. Previously only related strains were observed in Bolivian schoolchildren with the same serotype within a colonisation event (Inverarity et al., 2010). These data are important as there are few detailed studies addressing multiple pneumococcal carriage. To our knowledge, only two reports (Sá-Leão et al., 2002; Vestrheim et al., 2008) included serotyping, antibiotyping and genotyping data. However, they contained either a very small sample size (Sá-Leão et al., 2002) or had little information on colonisation of multiple strains of the same serotype (Vestrheim et al., 2008).

Other multiple colony carriage studies were limited to serotyping by antibody-based methods and used different culture methods (Hansman et al., 1985; Gratten et al., 1989; Huebner et al., 2000; Meats et al., 2003; Bronsdon et al., 2004; O'Brien et al., 2007; Hare et al., 2008; Hill et al., 2008; Kaltoft et al., 2008; Auranen et al., 2010). These studies showed large variation of co-colonisation rates with a low number of serotypes present together. While our method detected a multiple colonisation prevalence that is within the range published by others, we have detected – and confirmed by DNA-based methods – more serotypes colonising together with further diversity by sequence type and antibiotype. In the seminal work by Gundel and Okura (1933), all the colonisation events studied had multiple serotypes with a maximum of five serotypes. Although we found a much lower level of multiple serotype colonisations (11.5%) we found a similar maximum diversity of serotypes. Gundel and Okura's methodology involved passaging nasopharyngeal samples through mice repeatedly, which is laborious, time-consuming, and expensive. Random colony selection for serotyping was used in this study to better represent multiple carriage compared to selection by colony morphology (Hare et al., 2008). Mathematical derivations showed that increasing the colony number can uncover serotypes rarely detected in the nasopharynx (Huebner et al., 2000). This derivation indicates that our method would reveal serotypes that constitute 14% of the nasopharynx with 95% confidence. However, a correlation between the number of colonies characterised and the prevalence of multiple strain colonisations cannot be made from these studies as different methodologies were used. The cohort studied and its geographical location are also likely to influence the prevalence of co-colonisation and the extent of phenotype and genotype diversity. Although serotyping of colony sweeps and DNA-based methods such as microarrays can further enhance the sensitivity of detecting multiple serotypes, detection of rare serotypes is difficult with colony sweep serotyping, and microarray detection is limited to a handful of reference laboratories due to the high cost (Donkor et al., 2011; Turner et al., 2011). Additionally, microarray cannot differentiate between viable and non-viable bacteria.

This cohort of children is part of a semi-isolated community residing on a sugar plantation without a pneumococcal vaccination programme. Thus, our observations provide an insight into the diversity of pneumococcal colonisation in the absence of pneumococcal immunization. A recent study has analysed the effects of pneumococcal vaccination on the carriage of multiple serotypes (Brugger et al., 2010). While the prevalence of multiple serotype carriage did not alter 3 years after the use of heptavalent conjugate vaccine (PCV7), the prevalence of rare serotypes increased. Thus, should this area be subjected to future vaccination schemes, this data would be critical baseline data for studying the impact of vaccination on the ecology of multiple carriage.

An important observation in this study was that 57% (4/7) of multiple colonisation episodes had mixed antibiotypes illustrating the potential for genetic exchange between strains with different drug susceptibility. For example, in four colonisation episodes there were both penicillin susceptible and non-susceptible strains colonising together. The presence of penicillin susceptible together with non-susceptible strains has previously been reported in two other studies (Gratten et al., 1989; Sá-Leão et al., 2002). Two episodes also had simultaneous colonisation of Sxt susceptible and non-susceptible pneumococci, and this has not been previously reported. The spread of penicillin resistance in pneumococci has been attributed to the HGT of *pbp2x*, *pbp2b*, and *pbp1a*, predominantly through mosaic blocks (Dowson et al., 1989; Laible and Hakenbeck, 1991; Hakenbeck et al., 1998). The presence of mosaics resembling *S. mitis* sequences has been associated with a penicillin MIC of over 16 μg/mL (Soriano et al., 2008). Similarly, despite a single nucleotide mutation conferring co-trimoxazole resistance (Adrian and Klugman, 1997), HGT between pneumococci and viridans streptococci have been proposed to contribute to non-susceptibility to Sxt (Wilen et al., 2009). Therefore, a mixed colonisation of antibiotic-susceptible and non-susceptible strains increases the potential for susceptible strains to acquire resistant determinants, enabling them to survive the presence of antibiotic.

Sequence typing (ST) revealed that 58% were novel, 26% were previously only isolated from Africa, and 16% were previously isolated outside Africa. Geographically unique STs were also found in Latin America (Reis et al., 2008; Inverarity et al., 2011), and the Gambia (Donkor et al., 2011). It has been postulated that there is a greater potential for new STs to arise if the carriage rate of multiple strains is higher (Donkor et al., 2011). Although, in this study, the carriage of multiple strains was 12%, lower than the 19% reported by Donkor et al. (2011) the diversity of strains within the multiple colonisation events of our cohort was greater, which would also increase the potential for generating new STs. Despite the serotype diversity observed in multiple colonisations our sequence typing data did not reveal any evidence of capsule switching. Capsule switching rates are unclear; some studies reported that recombination of *cps* loci to be relatively rare (Brueggemann et al., 2007; Sandgren et al., 2004), while others indicate a high rate of switching (Jefferies et al., 2004; Stanhope et al., 2007).

eBURST analysis of STs, which is based on a clonal expansion and diversification model (Feil et al., 2004) revealed the presence of genetically related SLVs and DLVs strains circulating in our cohort, with each pair of related strains having the same serotype. Although eBURST analysis only links DLVs to a clonal complex if an intermediate SLV is present they could still be related. The majority of the genetically related strains with the same serotype differed in their variant loci by multiple nucleotides, suggesting that diversification, or "microevolution" has arisen through HGT events rather than multiple genetic mutations (Feil et al., 2000). HGT has recently been shown to contribute to the microevolution of pneumococci even within a single lineage, where extensive recombination occurred in genes encoding antibiotic targets, capsules, and other surface proteins (Croucher et al., 2011). Two pairs of *ddl* alleles differed in sequence by almost 10% with allele 108 showing a high sequence similarity to that of viridans streptococci, suggesting inter-species HGT has occurred. While genetically related strains are more likely to have arisen through evolution, the possibility of them being separate lineages cannot be excluded. Thus we are planning to perform whole genome sequencing on these strains.

Furthermore, sequence typing revealed that the same strain (ST4432) of a 6B serotype was isolated 5 months later in Child 11, and a new strain of serotype 6B was isolated in Child 3 6 months later (ST4368 then ST4373). However, the absence of capturing pneumococci in the interceding months that indicates re-acquisition is inconclusive as these strains may persist below the level of detection. From an immunological perspective, our data on re-acquisition of a different strain with the same serotype may provide further support that colonisation does not provide complete immunity, and that protective immunity to colonisation is not solely dependent on capsular polysaccharide (McCool et al., 2002). This is borne out by the absence of a correlation between a previous colonisation by serotype 6B and the titre of serotype 6B specific antibodies (Weinberger et al., 2008). Thus, additional surface proteins may elicit protective immunity to colonisation, and these may be accessory genes that are not present in all strains of a particular serotype, or within a clonal complex of related STs, or even within a particular ST (Donati et al., 2010).

5. Conclusion

Our method of randomly selecting and characterising up to 20 colonies per episode revealed a high number of serotypes and antibiotic-susceptible and non-susceptible strains colonising together. Further strain diversity was observed by sequence typing. We have shown that serotyping alone underestimated the diversity of pneumococci co-colonising the nasopharynx.

Acknowledgments

We are grateful to the directors, staff, parents, guardians, and children of the TPC Hospital. We acknowledge the use of the pneumococcal MLST database (Imperial College London) funded by the Wellcome Trust. We thank Dr. C. Bishop, the curator of the pneumococcal MLST database and Dr. C. Ling for independent verification of the sequence chromatograms. We acknowledge the support of the Special Trustees Charity of the Royal Free Hospital.

References

Adegbola, R.A., Obaro, S.K., Biney, E., Greenwood, B.M., 2001. Evaluation of Binax now *Streptococcus pneumoniae* urinary antigen test in children in a community with a high carriage rate of pneumococcus. Pediatr. Infect. Dis. J 20, 718–719.

Adrian, P.V., Klugman, K.P., 1997. Mutations in the dihydrofolate reductase gene of trimethoprim-resistant isolates of *Streptococcus pneumoniae*. Antimicrob. Agents Chemother. 41, 2406–2413.

Auranen, K., Mehtala, J., Tanskanen, A., Kaltoft, M.S., 2010. Between-strain competition in acquisition and clearance of pneumococcal carriage-epidemiologic evidence from a longitudinal study of day-care children. Am. J. Epidemiol. (171), 169–176.

Bentley, S.D., Aanensen, D.M., Mavroidi, A., Saunders, D., Rabbinowitsch, E., Collins, M., Donohoe, K., Harris, D., Murphy, L., Quail, M.A., Samuel, G., Skovsted, I.C., Kaltoft, M.S., Barrell, B., Reeves, R., Parkhill, J., Spratt, B.G., 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet. (2), e31.

Bogaert, D., De Groot, R., Hermans, P.W., 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. Lancet Infect. Dis. (4), 144–154.

Bronsdon, M.A., O'Brien, K.L., Facklam, R.R., Whitney, C.G., Schwartz, B., Carlone, G.M., 2004. Immunoblot method to detect *Streptococcus pneumoniae* and identify multiple serotypes from nasopharyngeal secretions. J. Clin. Microbiol. (42), 1596–1600.

Brueggemann, A.B., Pai, R., Crook, D.W., Beall, B., 2007. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. PLoS Pathog. (3), e168.

Brugger, S.D., Frey, P., Aebi, S., Hinds, J., Mühlemann, K., 2010. Multiple colonization with *S. pneumoniae* before and after introduction of the seven-valent conjugated pneumococcal polysaccharide vaccine.. PLoS ONE (5), e11638.

Charalambous, B.M., Oriyo, N.M., Gillespie, S.H., 2008. How valid is single-colony isolation for surveillance of *Streptococcus pneumoniae* carriage? J. Clin. Microbiol. (46), 2467–2468.

Clinical Laboratory and Standards Institute, 2010. Performance Standards for Antimicrobial Susceptibility Testing; Twentieth Informational Supplement. (88).

Coffey, T.J., Enright, M.C., Daniels, M., Morona, J.K., Morona, R., Hryniewicz, W., Paton, J.C., Spratt, B.G., 1998. Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of *Streptococcus pneumoniae*. Mol. Microbiol. (27), 73–83.

Croucher, N.J., Harris, S.R., Fraser, C., Quail, M.A., Burton, J., van der Linden, M., McGee, L., von Gottberg, A., Song, J.H., Ko, K.S., Pichon, B., Baker, S., Parry, C.M., Lambertsen, L.M., Shahinas, D., Pillai, D.R., Mitchell, T.J., Dougan, G., Tomasz, A., Klugman, K.P., Parkhill, J., Hanage, W.P., Bentley, S.D., 2011. Rapid pneumococcal evolution in response to clinical interventions. Science 331, 430–434.

Dagerhamn, J., Blomberg, C., Browall, S., Sjostrom, K., Morfeldt, E., Henriques-Normark, B., 2008. Determination of accessory gene patterns predicts the same relatedness among strains of *Streptococcus pneumoniae* as sequencing of housekeeping genes does and represents a novel approach in molecular epidemiology. J. Clin. Microbiol. 46, 863–868.

Donati, C., Hiller, N.L., Tettelin, H., Muzzi, A., Croucher, N.J., Angiuoli, S.V., Oggioni, M., Dunning Hotopp, J.C., Hu, F.Z., Riley, A., Covacci, D.R., Mitchell, T.J., Bentley, S.D., Kilian, M., Ehrlich, G.D., Rappuoli, R., Moxon, E.R., Masignani, V., 2010. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. Genome Biol. (11), R107.

Donkor, E.S., Bishop, C.J., Antonio, M., Wren, B., Hanage, W.P., 2011. High levels of recombination among *Streptococcus pneumoniae* isolates from the Gambia. MBio (2).

Dowson, C.G., Coffey, T.J., Kell, C., Whiley, R.A., 1993. Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of *Streptococcus mitis* in the formation of a low affinity PBP2B in *S. Pneumoniae*. Mol. Microbiol. 9, 635–643.

Dowson, C.G., Hutchison, A., Brannigan, J.A., George, R.C., Hansman, D., Linares, J., Tomasz, A., Smith, J.M., Spratt, B.G., 1989. Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. Proc. Natl. Acad. Sci. USA 86, 8842–8846.

Enright, M.C., Spratt, B.G., 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology 144 (Pt 11), 3049–3060.

Feil, E.J., Enright, M.C., Spratt, B.G., 2000. Estimating the relative contributions of mutation and recombination to clonal diversification: a comparison between Neisseria meningitidis and *Streptococcus pneumoniae*. Res. Microbiol. 151, 465–469.

Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., Spratt, B.G., 2004. EBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. 186, 1518–1530.

Gratten, M., Montgomery, J., Gerega, G., Gratten, H., Siwi, H., Poli, A., Koki, G., 1989. Multiple colonization of the upper respiratory tract of Papua New Guinea children with Haemophilus influenzae and *Streptococcus pneumoniae*. Southeast Asian J. Trop. Med. Public Health 20, 501–509.

Gundel, M., Okura, G., 1933. Untersuchungen über das gleichzeitige Vorkommen mehrerer Pneumokokkentypen bei Gesunden und ihre Bedeutung für die Epidemiologie. Med. Microbiol. Immunol. 114, 678–704.

Hakenbeck, R., Konig, A., Kern, I., van der Linden, M., Keck, W., Billot-Klein, D., Legrand, R., Schoot, B., Gutmann, L., 1998. Acquisition of five high-Mr penicillin-binding protein variants during transfer of high-level beta-lactam resistance from *Streptococcus mitis* to *Streptococcus pneumoniae*. J. Bacteriol. 180, 1831–1840.

Hansman, D., Morris, S., Gregory, M., McDonald, B., 1985. Pneumococcal carriage amongst Australian aborigines in Alice Springs, Northern Territory. J. Hyg. (Lond.) 95, 677–684.

Hare, K.M., Morris, P., Smith-Vaughan, H., Leach, A.J., 2008. Random colony selection versus colony morphology for detection of multiple pneumococcal serotypes in nasopharyngeal swabs. Pediatr. Infect Dis. J. 27, 178–180.

Hill, P.C., Cheung, Y.B., Akisanya, A., Sankareh, K., Lahai, G., Greenwood, B.M., Adegbola, R.A., 2008. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. Clin. Infect. Dis. 46, 807–814.

Hiller, N.L., Ahmed, A., Powell, E., Martin, D.P., Eutsey, R., Earl, J., Janto, B., Boissy, R.J., Hogg, J., Barbadora, K., Sampath, R., Lonergan, S., Post, J.C., Hu, F.Z., Ehrlich, G.D., 2010. Generation of genic diversity among *Streptococcus pneumoniae* strains via horizontal gene transfer during a chronic polyclonal pediatric infection. PLoS Pathog. 6, 1–14.

Hiller, N.L., Janto, B., Hogg, J.S., Boissy, R., Yu, S., Powell, E., Keefe, R., Ehrlich, N.E., Shen, K., Hayes, J., Barbadora, K., Klimke, W., Dernovoy, D., Tatusova, T., Parkhill, J., Bentley, S.D., Post, J.C., Ehrlich, G.D., Hu, F.Z., 2007. Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome. J. Bacteriol. 189, 8186–8195.

Hogg, J.S., Hu, F.Z., Janto, B., Boissy, R., Hayes, J., Keefe, R., Post, J.C., Ehrlich, G.D., 2007. Characterization and modeling of the Haemophilus influenzae core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains. Genome Biol. (8), R103.

Huebner, R.E., Dagan, R., Porath, N., Wasas, A.D., Klugman, K.P., 2000. Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. Pediatr. Infect Dis. J. 19, 1017–1020.

Huebner, R.E., Wasas, A., Mushi, A., Mazhani, L., Klugman, K., 1998. Nasopharyngeal carriage and antimicrobial resistance in isolates of *Streptococcus pneumoniae* and Haemophilus influenzae type b in children under 5 years of age in Botswana. Int. J. Infect. Dis. 3, 18–25.

Inverarity, D., Diggle, M., Ure, R., Johnson, P., Altstadt, P., Mitchell, T., Edwards, G., Santana-Hernandez, D., 2011. Molecular epidemiology and genetic diversity of pneumococcal carriage among children in Beni State. Bolivia. Trans. R. Soc. Trop. Med. Hyg. 105, 445–451.

Inverarity, D., Diggle, M., Ure, R., Santana-Hernandez, D., Altstadt, P., Mitchell, T., Edwards, G., 2010. Simultaneous nasopharyngeal carriage of two pneumococcal multilocus sequence types with a serotype 3 phenotype. Int. J. Microbiol., 765479.

Jefferies, J.M., Smith, A., Clarke, S.C., Dowson, C., Mitchell, T.J., 2004. Genetic analysis of diverse disease-causing pneumococci indicates high levels of diversity within serotypes and capsule switching. J. Clin. Microbiol. 42, 5681–5688.

Kaltoft, M.S., Skov Sorensen, U.B., Slotved, H.C., Konradsen, H.B., 2008. An easy method for detection of nasopharyngeal carriage of multiple *Streptococcus pneumoniae* serotypes. J. Microbiol. Methods 75, 540–544.

Kong, F., Wang, W., Tao, J., Wang, L., Wang, Q., Sabananthan, A., Gilbert, G.L., 2005. A molecular-capsular-type prediction system for 90 *Streptococcus pneumoniae* serotypes using partial cpsA–cpsB sequencing and wzy- or wzx-specific PCR. J. Med. Microbiol. 54, 351–356.

Laible, G., Hakenbeck, R., 1991. Five independent combinations of mutations can result in low-affinity penicillin-binding protein 2x of *Streptococcus pneumoniae*. J. Bacteriol. 173, 6986–6990.

Leung, M., Freystatter, K., Bryson, K., Pichon, B., Gillespie, S.H., Charalambous, B.M. A new method for determining *Streptococcus pneumoniae* serotype from a single PCR amplicon. in preparation.

Lipsitch, M., Dykes, J.K., Johnson, S.E., Ades, E.W., King, J., Briles, D.E., Carlone, G.M., 2000. Competition among *Streptococcus pneumoniae* for intranasal colonization in a mouse model. Vaccine 18, 2895–2901.

McCool, T.L., Cate, T.R., Moy, G., Weiser, J.N., 2002. The immune response to pneumococcal proteins during experimental human carriage. J. Exp. Med. 195, 359–365.

Meats, E., Brueggemann, A.B., Enright, M.C., Sleeman, K., Griffiths, D.T., Crook, D.W., Spratt, B.G., 2003. Stability of serotypes during nasopharyngeal carriage of *Streptococcus pneumoniae*. J. Clin. Microbiol. 41, 386–392.

O'Brien, K.L., Bronsdon, M.A., Dagan, R., Yagupsky, P., Janco, J., Elliott, J., Whitney, C.G., Yang, Y.H., Robinson, L.G., Schwartz, B., Carlone, G.M., 2001. Evaluation of a medium (STGG) for transport and optimal recovery of *Streptococcus pneumoniae* from nasopharyngeal secretions collected during field studies. J. Clin. Microbiol. 39, 1021–1024.

O'Brien, K.L., Millar, E.V., Zell, E.R., Bronsdon, M., Weatherholtz, R., Reid, R., Becenti, J., Kvamme, S., Whitney, C.G., Santosham, M., 2007. Effect of pneumococcal conjugate vaccine on nasopharyngeal colonization among immunized and unimmunized children in a community-randomized trial. J. Infect. Dis. 196, 1211–1220.

Pai, R., Gertz, R.E., Beall, B., 2006. Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. J. Clin. Microbiol. 44, 124–131.

Reis, J.N., Palma, T., Ribeiro, G.S., Pinheiro, R.M., Ribeiro, C.T., Cordeiro, S.M., da Silva Filho, H.P., Moschioni, M., Thompson, T.A., Spratt, B., Riley, L.W., Barocchi, M.A., Reis, M.G., Ko, A.I., 2008. Transmission of *Streptococcus pneumoniae* in an urban slum community. J. Infect. 57, 204–213.

Sá-Leão, R., Tomasz, A., Santos Sanches, I., de Lencastre, H., 2002. Pilot study of the genetic diversity of the pneumococcal nasopharyngeal flora among children attending day care centers. J. Clin. Microbiol. 40, 3577–3585.

Sandgren, A., Sjostrom, K., Olsson-Liljequist, B., Christensson, B., Samuelsson, A., Kronvall, G., Henriques Normark, B., 2004. Effect of clonal and serotype-specific properties on the invasive capacity of *Streptococcus pneumoniae*. J. Infect. Dis. 189, 785–796.

Shen, K., Antalis, P., Gladitz, J., Sayeed, S., Ahmed, A., Yu, S., Hayes, J., Johnson, S., Dice, B., Dopico, R., Keefe, R., Janto, B., Chong, W., Goodwin, J., Wadowsky, R.M., Erdos, G., Post, J.C., Ehrlich, G.D., Hu, F.Z., 2005. Identification, distribution, and expression of novel genes in 10 clinical isolates of nontypeable *Haemophilus influenzae*. Infect. Immun. 73, 3479–3491.

Shen, K., Gladitz, J., Antalis, P., Dice, B., Janto, B., Keefe, R., Hayes, J., Ahmed, A., Dopico, R., Ehrlich, N., Jocz, J., Kropp, L., Yu, S., Nistico, L., Greenberg, D.P., Barbadora, K., Preston, R.A., Post, J.C., Ehrlich, G.D., Hu, F.Z., 2006. Characterization, distribution, and expression of novel genes among eight clinical isolates of *Streptococcus pneumoniae*. Infect. Immun. 74, 321–330.

Sørensen, U.B., 1993. Typing of pneumococci by using 12 pooled antisera. J. Clin. Microbiol. 31, 2097–2100.

Soriano, F., Cafini, F., Aguilar, L., Tarrago, D., Alou, L., Gimenez, M.J., Gracia, M., Ponte, M.C., Leu, D., Pana, M., Letowska, I., Fenoll, A., 2008. Breakthrough in penicillin resistance? *Streptococcus pneumoniae* isolates with penicillin/cefotaxime MICs of 16 mg/L and their genotypic and geographical relatedness. J. Antimicrob. Chemother. 62, 1234–1240.

Stanhope, M.J., Walsh, S.L., Becker, J.A., Miller, L.A., Lefebure, T., Lang, P., Bitar, P.D., Amrine-Madsen, H., 2007. The relative frequency of intraspecific lateral gene transfer of penicillin binding proteins 1a, 2b, and 2x, in amoxicillin resistant *Streptococcus pneumoniae*. Infect. Genet. Evol. 7, 520–534.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., Deboy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Telford, M.R., Rappuoli, R., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. USA 102, 13950–13955.

Turner, P., Hinds, J., Turner, C., Jankhot, A., Gould, K., Bentley, S., Nosten, F., Goldblatt, D., 2011. Improved detection of nasopharyngeal co-colonization by multiple pneumococcal serotypes using latex agglutination or molecular serotyping by microarray. J. Clin. Microbiol. 49, 1784–1789.

Vestrheim, D.F., Hoiby, E.A., Aaberge, I.S., Caugant, D.A., 2008. Phenotypic and genotypic characterization of *Streptococcus pneumoniae* strains colonizing children attending day-care centers in Norway. J. Clin. Microbiol. 46, 2508–2518.

Weinberger, D.M., Dagan, R., Givon-Lavi, N., Regev-Yochay, G., Malley, R., Lipsitch, M., 2008. Epidemiologic evidence for serotype-specific acquired immunity to pneumococcal carriage. J. Infect. Dis. 197, 1511–1518.

Wilen, M., Buwembo, W., Sendagire, H., Kironde, F., Swedberg, G., 2009. Cotrimoxazole resistance of *Streptococcus pneumoniae* and commensal streptococci from Kampala. Uganda. Scand. J. Infect. Dis. 41, 113–121.

World Health Organization, 2007. Pneumococcal conjugate vaccine for childhood immunization – WHO position paper. Wkly. Epidemiol. Rec. 82, 93–104.

JCM
Journals.ASM.org

Streptococcus pseudopneumoniae Identification by Pherotype: a Method To Assist Understanding of a Potentially Emerging or Overlooked Pathogen

Marcus H. Leung,[a] Clare L. Ling,[a]* Holly Ciesielczuk,[a]* Julianne Lockwood,[a]* Sarah Thurston,[b] Bambos M. Charalambous,[a] and Stephen H. Gillespie[a]*

Centre for Clinical Microbiology, University College London Medical School, Royal Free Campus, London, United Kingdom,[a] and Department of Respiratory Medicine, University College London Medical School, Royal Free Campus, London, United Kingdom[b]

The recent identification of *Streptococcus pseudopneumoniae* (pseudopneumococcus) has complicated classification schemes within members of the "mitis" streptococcal group. Accurate differentiation of this species is necessary for understanding its disease potential and identification in clinical settings. This work described the use of the competence-stimulatory peptide ComC sequence for identification of *S. pseudopneumoniae*. ComC sequences from clinical sources were determined for 17 strains of *S. pseudopneumoniae*, *Streptococcus pneumoniae*, and *Streptococcus oralis*. An additional 58 ComC sequences from a range of sources were included to understand the diversity and suitability of this protein as a diagnostic marker for species identification. We identified three pherotypes for this species, delineated CSP6.1 (10/14, 79%), CSP6.3 (3/14, 21%), and SK674 (1/14, 7%). Pseudopneumococcal ComC sequences formed a discrete cluster within those of other oral streptococci. This suggests that the *comC* sequence could be used to identify *S. pseudopneumoniae*, thus simplifying the study of the pathogenic potential of this organism. To avoid confusion between pneumococcal and pseudopneumococcal pherotypes, we have renamed the competence pherotype CSP6.1, formerly reported as an "atypical" pneumococcus, CSPps1 to reflect its occurrence in *S. pseudopneumoniae*.

The mitis group of streptococci includes nasopharyngeal colonizers such as *Streptococcus mitis*, *Streptococcus oralis*, *Streptococcus pneumoniae*, and the recently classified *Streptococcus pseudopneumoniae* (2). Of these, *S. pneumoniae* (pneumococcus) is responsible for more than a million deaths annually and is responsible for diseases such as otitis media, pneumonia, septicemia, and meningitis. However, invasive diseases caused by other related viridans group streptococci had been documented (16, 26, 31).

Some strains of *S. pseudopneumoniae*, along with *S. mitis* and *S. oralis*, have often been classified previously as "atypical pneumococci," because of their similarity to *S. pneumoniae*. These organisms share ≥99% identity in 16S rRNA gene sequences (2, 21, 43). Optochin sensitivity and bile solubility, the two standard pneumococcal phenotypic identification tests, have proven to be inconclusive for differentiating pneumococci from these atypical strains (3, 4, 9, 10, 17, 19, 20, 22, 27, 29, 32, 34, 38, 39, 49). Virulence factors that were once thought to be exclusive to the pneumococcus, such as pneumolysin (encoded by *ply*) and autolysin A (encoded by *lytA*), have been detected in commensal streptococcal species (18, 35, 49), compromising their specificity as species identification markers. The pathogenic potential of *S. pseudopneumoniae* (the pseudopneumococcus) has been demonstrated in a murine model (12) as well as in humans (2, 18, 23, 24, 28, 40). Rapid, correct identification of this organism in the clinical setting is essential for diagnosis and for understanding its disease potential. A simple, unequivocal method to identify *S. pseudopneumoniae* would be valuable.

Streptococci are competent for genetic transformation. In the case of *S. pneumoniae*, this is mediated by the competence-stimulatory peptide (CSP) encoded by the *comC* gene (13). CSP sequences differ between species and within species; different versions within species are known as pherotypes (48). We report the distribution of the *comC* sequence in strains of *S. pseudopneumoniae* and show that it may prove a valuable method to identify the organism rapidly.

MATERIALS AND METHODS

Clinical specimens and bacteria. A total of 17 clinical specimens of *S. pseudopneumoniae*, *S. pneumoniae*, and *S. oralis* were collected at the Royal Free Hospital Microbiology Laboratory between the years 1993 and 2010 (Table 1). Sixteen samples were from patients with lower respiratory tract (LRT) infections, and a single strain was isolated from a normally sterile site. Samples were plated on Columbia blood agar (Oxoid, Cambridgeshire, United Kingdom) in 5% CO_2 at 35°C in an attempt to cultivate bacteria, and colonies suggestive of pneumococci based on morphology and alpha-hemolysis were tested for optochin sensitivity.

Genomic DNA extraction for amplification. Genomic DNA from LRT samples was extracted by a modified Chelex method as described previously (47). The supernatant containing the DNA was used as the amplification template. For culture-positive clinical specimens, genomic

Received 13 January 2012 Returned for modification 11 February 2012
Accepted 23 February 2012
Published ahead of print 29 February 2012

Address correspondence to Bambos M. Charalambous, b.charalambous@ucl.ac.uk.

* Present address: Clare L. Ling, Microbiology Services, Health Protection Agency, Barts and the London NHS Trust, London, United Kingdom; Holly Ciesielczuk, Centre for Immunology and Infectious Diseases, Blizard Centre, Queen Mary University of London, London, United Kingdom; Julianne Lockwood, Department of Virology, Great Ormond Street Hospital for Children, London, United Kingdom; Stephen H. Gillespie, School of Medicine, University of St. Andrews, North Haugh, St. Andrews, Fife, United Kingdom.

Supplemental material for this article may be found at http://jcm.asm.org/.

doi:10.1128/JCM.00131-12

TABLE 1 Streptococcal strains included for this study

Species	Strain	Clinical isolation site	Accession no.[c]	Source or reference
S. pseudopneumoniae	N452	Blood	Not deposited	This study
	RFH504	LRT[a]	Not deposited	This study
	RFH543	LRT	Not deposited	This study
	RFH686	LRT	Not deposited	This study
	RFH687	LRT	Not deposited	This study
	RFH827	LRT	Not deposited	This study
	RFH905	LRT	Not deposited	This study
	RFH999	LRT	Not deposited	This study
	874	Unknown[b]	AJ240773	48
	ATCC BAA-960	LRT	Not deposited	This study
	IS7493	LRT	YP004769537	40
	PT5479	Naso/oropharynx	Not deposited	41
	PT5779	Naso/oropharynx	Not deposited	41
	SK674	Unknown	Not deposited	25
S. pneumoniae	RFH324	LRT	Not deposited	This study
	RFH410	LRT	Not deposited	This study
	RFH577	LRT	Not deposited	This study
	RFH815	LRT	Not deposited	This study
	RFH864	LRT	Not deposited	This study
	RFH904	LRT	Not deposited	This study
	VA1	Unknown	AJ240789	48
	41G	Unknown	AJ240766	48
	CSP2.1b	Nasopharynx	Not deposited	46
	Pn24	Unknown	AJ240759	48
	Pn59	Unknown	AJ240793	48
	Pn13	Unknown	AJ240792	48
	101/87	Unknown	AJ240791	48
	SK676	Unknown	Not deposited	25
S. mitis	Col15	Unknown	AJ240762	48
	Col16	Unknown	AJ240763	48
	NCTC 10712	LRT	AJ240795	48
	NCTC 12261	Naso/oropharynx	AJ000875	15
	B5	Unknown	AJ000871	15
	B6	Unknown	AJ000865	15
	Hu8	Unknown	AJ000866	15
	SK137	Unknown	Not deposited	25
	SK145	Unknown	Not deposited	25
	SK262	Unknown	Not deposited	25
	SK272	Unknown	Not deposited	25
	SK564	Unknown	Not deposited	25
	SK596	Unknown	Not deposited	25
	SK598	Unknown	Not deposited	25
	SK599	Unknown	Not deposited	25
	SK601	Unknown	Not deposited	25
	SK602	Unknown	Not deposited	25
	SK608	Unknown	Not deposited	25
	SK609	Unknown	Not deposited	25
	SK611	Unknown	Not deposited	25
	SK612	Unknown	Not deposited	25
	SK614	Unknown	Not deposited	25
	SK615	Unknown	Not deposited	25
	SK667	Unknown	Not deposited	25
	SK675	Unknown	Not deposited	25
S. oralis	RFH623	LRT	Not deposited	This study
	RFH831	LRT	Not deposited	This study
	Col19	Unknown	AJ240794	48
	NCTC 11427	Naso/oropharynx	AJ000873	15
	DSM 20066	Unknown	AJ000874	15

(Continued on following page)

TABLE 1 (Continued)

Species	Strain	Clinical isolation site	Accession no.[c]	Source or reference
	SK153	Unknown	Not deposited	25
	SK305	Unknown	Not deposited	25
	SK34	Unknown	Not deposited	25
	SK39	Unknown	Not deposited	25
	SK571	Unknown	Not deposited	25
	SK597	Unknown	Not deposited	25
	SK610	Unknown	Not deposited	25
	SK92	Unknown	Not deposited	25
S. gordonii	NCTC 3165	Gum	AJ000870	15
	NCTC 7865	Endocardium	X98110	14
	NCTC 7868	Unknown	X98109	14
S. infantis	SK140	Unknown	Not deposited	25
	SK282	Unknown	Not deposited	25
	SK283	Unknown	Not deposited	25
	SK350	Unknown	Not deposited	25
S. cristatus	NCTC 12479	Unknown	AJ000876	15
S. peroris	ATCC 700780	Tooth	EFX39822	NCBI genome

[a] LRT, lower respiratory tract.
[b] Unknown, isolation site not specified in previous publications or not available on ATCC or NCTC database.
[c] Accession numbers are absent where strain sequences were under 200 bp and thus not deposited. Nucleotide sequences for pherotypes associated with strains characterized in this study are given in Table S2 in the supplemental material (pherotype CSPps1a is associated with *S. pseudopneumoniae* strains N452, RFH504, RFH543, RFH687, RFH905, and RFH999; pherotype CSPps2b is associated with *S. pseudopneumoniae* strains RFH686, RFH827, and ATCC BAA-960; pherotype CSP1c is associated with *S. pneumoniae* strains RFH324, RFH410, RFH577, RFH815, RFH864, and RFH904; and pherotype CSP6.2d is associated with *S. oralis* strains RFH623 and RFH831). Undeposited sequences for other strains listed are available in the indicated reference.

DNA was extracted using the Wizard genomic DNA purification kit (Promega) or by the heat lysis method (30).

Presumptive identification of *S. pseudopneumoniae*. Quantitative PCR (qPCR) using primers and probes specific for Spn9802 (1, 44) and *lytA* (5) was performed sequentially to differentiate between *S. pneumoniae* (both positive) and *S. pseudopneumoniae* (Spn9802 positive, *lytA* negative). To monitor PCR inhibition, a SPUD potato gene internal amplification control (IAC) was included in each reaction using primers targeting *phyB* of *Solanum tuberosum* (36). For *lytA* and Spn9802 qPCR assays, amplification reactions using 25-μl mixtures containing 1× Platinum quantitative PCR SuperMix-UDG (Invitrogen), a final concentration of 7 mM $MgCl_2$, primers and probes, 4×10^{-7} μM IAC template DNA (Sigma-Aldrich), and 5 μl of template DNA (see Table S1 in the supplemental material) were performed. Negative and positive controls were performed for each qPCR assay. Amplification was performed using a Rotor-Gene Q (Qiagen) with the following conditions: an initial hold cycle at 95°C for 3 min, followed by 40 cycles of 95°C for 15 s and 60°C for 45 s. The PCR data were acquired at the end of each cycle and analyzed by the instrument software (Qiagen). Samples with cycle threshold (C_T) values ≤35 for the *lytA* or Spn9802 target were considered positive; samples that had no C_T value for either the *lytA* or Spn9802 target but that did have a C_T value ≤40 for the IAC target were considered negative, and samples with a no C_T value for either the *lytA* or Spn9802 target and no C_T value for the IAC target were considered inhibited. Samples positive for both *lytA* and Spn9802 were considered to be *S. pneumoniae* positives, and samples positive for Spn9802 and negative for *lytA* were considered to be presumptively *S. pseudopneumoniae* positive. All reactions were performed in triplicate.

Amplification of *comC*. Forward and reverse primers were designed to bind *comC* at positions 592 to 611 and 873 to 892, respectively (accession number U33315). The primer sequences and primer reaction concentrations are indicated in Table S1 in the supplemental material. For strains grown on agar, genomic DNA was extracted from bacterial cells from a fresh overnight culture on Columbia blood agar by the heat lysis method as described previously (30). In addition to primers, each reaction mixture contained 2 μl of template DNA, 0.3 μl *Taq* polymerase (5 U/μl) (Invitrogen), 1× PCR buffer (10×), 3 μM $MgCl_2$, and 0.6 mM deoxynucleoside triphosphates (dNTPs; Promega), made up to a final volume of 50 μl with DNase/RNase-free distilled water (Gibco).

MLST. The primers used for multilocus sequence typing (MLST) have been described previously (8). MLST was performed with the same reaction components and concentrations indicated for the *comC* amplification described above.

Amplicon purification and sequencing of *comC* and MLST loci. Amplicons were analyzed with 1.5% (wt/vol) agarose gel electrophoresis. Amplicons with the expected band sizes were purified using a PCR purification kit (Qiagen) according to the manufacturer's instructions. Cycle sequencing and sequence analysis were performed in the same manner for both PCRs (see below). Purified DNA was sequenced using BigDye Sequencing Terminator v.3.1 (Applied Biosystems), and sequences were analyzed on the 3130 genetic analyzer (Applied Biosystems) and viewed on Bionumerics software (version 5).

Publicly accessible sequences of streptococcal CSP. Additional streptococcal ComC amino acid sequences were included in this study for the construction of a phylogenetic tree (see below). Sequences obtained from previous publications, with accession numbers, are indicated in Table 1. Sequences for which no accession numbers are listed have not been deposited in GenBank. The sequences can be found in the sources given in Table 1.

Construction of a *comC* phylogenetic tree. Multiple alignment of *comC* sequences was constructed with ClustalW functionality on MEGA version 5.05 (45). For both analyses, neighbor-joining trees were constructed based on alignment data. Bootstrap support of 1,000 repetitions was performed.

RESULTS

***S. pseudopneumoniae* CSP sequences.** It was possible to identify ComC sequences for nine pseudopneumococcal strains where

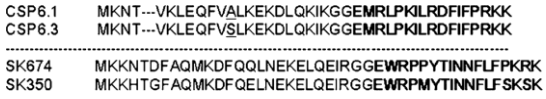


FIG 1 Pherotypes of *S. pseudopneumoniae* CSP6.1, CSP6.3, and SK674. The mature region of the peptide is in boldface after the double glycine. CSP6.1 and CSP6.3 differ by a single amino acid at position 12 (alanine in CSP6.1 and serine in CSP6.3; underlined). SK674 is a presumptive pseudopneumococcus (25), with an extended ComC compared to CSP6.1 and CSP6.3, and shares higher identity to *S. infantis* SK350. A total of three pseudopneumococcal pherotypes were characterized.

CSP sequences were derived directly from samples submitted to our laboratory from patients with LRT and invasive infections. These samples were presumed to contain pseudopneumococci based on either Spn9802-positive and *lytA*-negative qPCRs (eight strains), bile insolubility and optochin-variable phenotypic traits (one strain, N452), or MLST (one strain, N452). The pherotype of the control strain, ATCC BAA-960, was also characterized in this study. Two pherotypes were detected in these nine strains, six (N452, RFH504, RFH543, RFH687, RFH905, RFH999) of which were associated with CSP6.1. To avoid confusion with pherotypes of other oral streptococci, we propose that CSP6.1 be named CSPps1, where "ps" represents the pseudopneumococcus. The three remaining strains, BAA-960, RFH686, and RFH827, had a ComC sequence that has not been reported before; this sequence is identical to that of CSPps1 in size, differing by an alanine-to-serine substitution at position 12 of the propeptide. We propose that this pherotype be classified as CSPps2 (Fig. 1).

CSP sequences of four strains (IS7493, PT5479, PT5779, and SK674) were available from reports in earlier publications (25, 41, 48). All but that of SK674 were identical to CSPps1. Thus, more than 70% of the presumptive *S. pseudopneumoniae* strains in this study were associated with CSPps1 (Table 2). SK674 has an extended ComC of 54 amino acids, most similar to a pherotype characterized in *Streptococcus infantis* SK350, with six amino acid substitutions, three of which are in the mature peptide region (25) (Fig. 1 and Table 2).

Phylogenetic analysis of streptococcal pherotypes. A phylogenetic tree constructed from alignment of streptococcal pherotypes shows that, by pherotype, streptococci fall into two major groups (Fig. 2). *S. pneumoniae*, *S. pseudopneumoniae*, *S. mitis*, and some *S. oralis* strains belong to one group (group 1), while a more divergent and loosely defined group (group 2) consists of predominantly *Streptococcus gordonii*, *S. infantis*, *Streptococcus peroris*, *Streptococcus cristatus*, and most of the remaining *S. oralis* strains. SK674, formerly classified as a pseudopneumococcus (25), clustered near members of *S. infantis* by ComC alignment. All of the remaining *S. pseudopneumoniae* pherotypes are grouped in a separate cluster in close relation to other species, notably *S. oralis* and *S. mitis* (Fig. 2).

DISCUSSION

We have characterized pherotypes associated with *S. pseudopneumoniae* by comparing strains available to us, amplifying sequences collected from lower respiratory tract samples and collecting publicly available ComC sequences for this organism and related streptococcal species. We have shown that CSP6.1 (or CSPps1) is the commonest pherotype among *S. pseudopneumoniae* strains found in different geographical regions. Pherotype CSP6.1 was previously considered to be a rare pherotype of an "atypical nontypeable pneumococcal" strain, 874, based on multilocus sequence analysis (33, 48). Its classification as a pneumococcus may stem from its possession of *ply* and *lytA*, which were once considered suitable genetic markers for this organism (33). However, it is known that these two genes are not specific to *S. pneumoniae* (18, 35, 49), and we have been unable to find a report of CSP6.1 being found in a strain unequivocally identified as *S. pneumoniae*. Based on these observations, we hypothesize that strain 874 is a strain of *S. pseudopneumoniae*. Pneumococcal strain 101/87, associated with CSP5, was described as an "atypical pneumococcus" and could not be serotyped by Whatmore et al. (48), and this strain is most likely to be *S. pseudopneumoniae*. Phylogenetic analysis of CSP sequences in this study suggests that CSP5 is most closely related to *S. mitis* or *S. oralis* pherotypes that cluster together. Thus, we believe that CSP6.1 is associated with *S. pseudopneumoniae* and that, to differentiate pseudopneumococcal pherotypes from ComC sequences from other organisms, they should be designated CSPps1 instead of CSP6.1.

In this study we have identified a new pherotype associated with *S. pseudopneumoniae* and have designated this CSPps2. *S. pseudopneumoniae* pherotypes form a distinct cluster within those of other oral streptococcal species, suggesting that pseudopneumococcal pherotypes could be species specific and could be used as a simple diagnostic tool.

In contrast, strain SK674, a *S. pseudopneumoniae* strain identified based on clustering of housekeeping gene sequences (25), clustered closely in ComC sequences of *S. infantis* strains. One might argue that SK674 acquired a divergent *comC* from *S. infantis* by horizontal gene transfer. While interspecies transfer of the competence operon has been documented (15), the lack of such an event in our collection of over 200 pneumococcal strains indicates that this is relatively rare (our unpublished data). SK674 has a genome size of 1.87 Mbp (25), comparable to those of other *S. infantis* strains (1.74 to 1.88 Mbp) and much smaller than the genome size of the pseudopneumococcal strain IS7493 (2.1 Mbp) (40). It seems that the alternative explanation, that SK674 is actually a strain of *S. infantis*, is more likely. Additional analysis of SK674, such as DNA-DNA hybridization and *lytA* sequence analysis, may shed light as to the true identity of this strain.

From the recently characterized genome sequence of *S. pseudopneumoniae* strain IS7493, it was concluded that CSP-mediated induction of fratricide does not take place based on the absence of *comC* in this strain (40). We were, however, able to locate the gene that is identical to that of CSPps1. Thus, we can conclude that this *S. pseudopneumoniae* strain does contain the necessary gene sequences for production of a competence peptide and that a homologous gene is found in all pseudopneumococcal strains analyzed thus far, strengthening the use of this gene as a diagnostic marker.

TABLE 2 Distribution of pseudopneumococcal pherotypes

Pherotype	No. of strains (% of 14)	Strain(s)
CSP 6.1 (CSPps1)	10 (71)	N452, RFH504, RFH543, RFH687, RFH905, RFH999, IS7493, 874, PT5479, PT5779
CSPps2	3 (21)	BAA-960, RFH686, RFH827
SK674	1 (7)	SK674

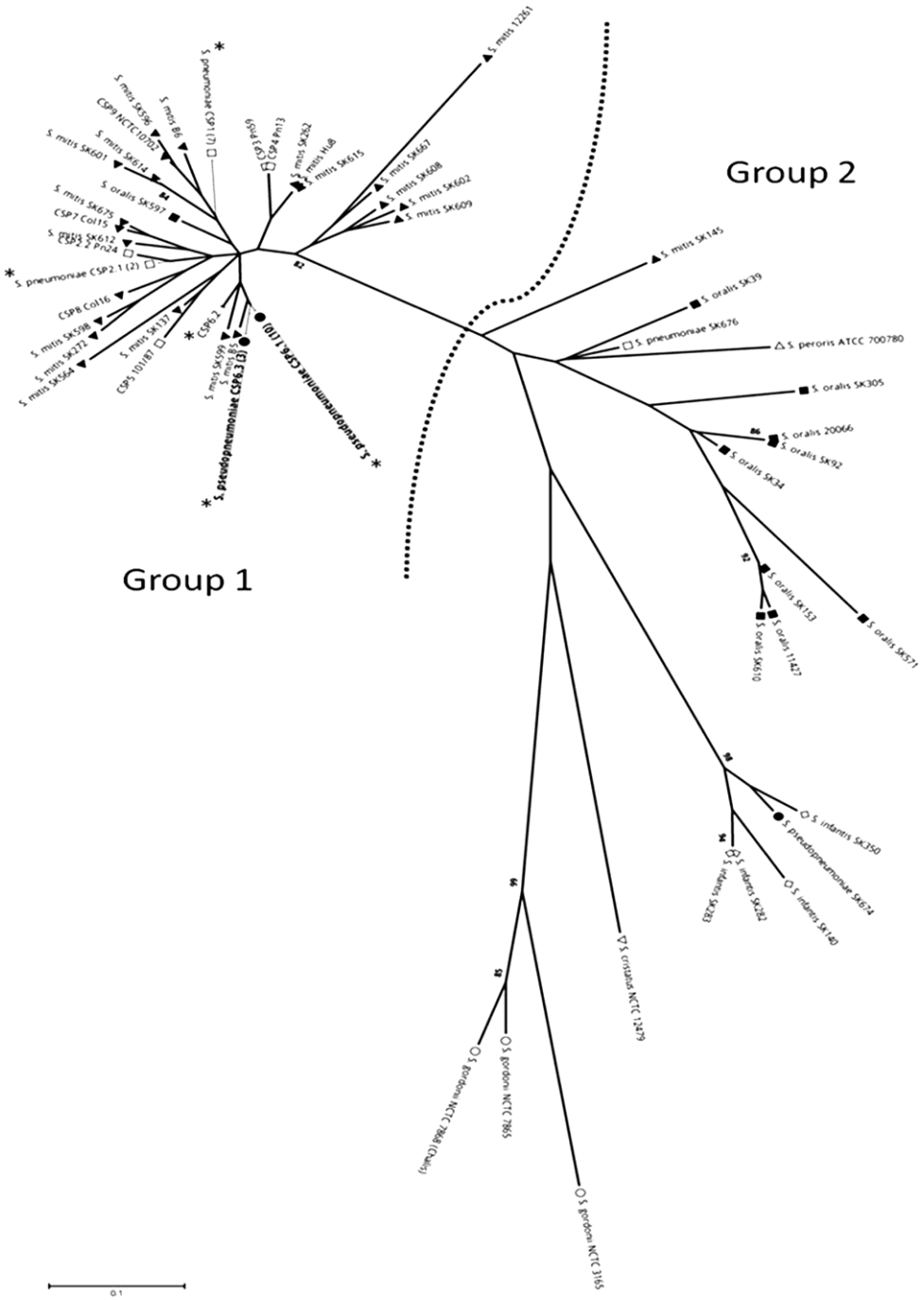


FIG 2 Neighbor-joining phylogenetic tree of ComC amino acid sequences of streptococcal species of the mitis group. Two clusters of ComC based on amino acid sequence similarities (groups 1 and 2) are separated by a dotted line. Each pherotype is labeled according to the classified species: filled circle, *S. pseudopneumoniae*; open square, *S. pneumoniae*; filled triangle, *S. mitis*; filled square, *S. oralis*; open circle, *S. gordonii*; open upright triangle, *S. peroris*; open diamond, *S. infantis*, open inverted triangle, *S. cristatus*. Asterisks indicate multiple strains of the same pherotype that have been compressed for clarity. Numbers of strains with the same pherotype are indicated in parentheses. Pseudopneumococcal strains of these groups are indicated in Table 2. The pneumococcal CSP1 pherotype includes strains VA1, RFH324, RFH410, RFH577, RFH815, RFH864, and RFH904. The pneumococcal CSP2.1 pherotype includes strains 41G and CSP2.1b. CSP6.2 pherotypes include SK671 (*S. mitis*), Col19 (*S. oralis*), RFH623 (*S. oralis*), and RFH831 (*S. oralis*). The phylogenetic tree was built with 1,000 bootstrap repetitions, with support over 80 indicated. The ruler indicates amino acid substitutions per site. The tree was constructed with MEGA 5.05.

S. pseudopneumoniae is usually identified as acapsulate, bile insoluble, and intermediately optochin resistant in 5% CO_2 and optochin susceptible in ambient O_2 (2). These tests can be difficult to standardize in the laboratory. Identification by *comC* sequencing would allow a rapid method of definitive diagnosis as the competence ligand gene is conserved across streptococcal species and the pseudopneumococcal *comC* sequences appear to provide taxonomic information, similar to the case of *gyr* in *Serratia* species (6). Previously reported genetic approaches to differentiate *S. pseudopneumoniae* from *S. pneumoniae* relied on targeting *lytA*, *cpsA*, *aliB*-like ORF2, *ply*, *psaA*, Spn9802, and *sodA* (2, 5, 7, 16, 24, 37, 42, 44); however, these are inconclusive. While multilocus sequence analysis can reveal divergence between species by their housekeeping and virulence gene fragments (11, 25), these methods are too cumbersome for large-scale and routine clinical diagnosis. Recent accounts of detection of pseudopneumococcus in carriage and symptomatic hosts with antibiotic resistance necessitate its accurate diagnosis as an emerging causative agent of disease (23). A recent report suggested that sequencing *recA* could differentiate between *S. pneumoniae* and *S. pseudopneumoniae* (43), but the study was performed with a smaller number of pseudopneumococcal strains solely from North America. Here we propose that pherotyping may be a promising diagnostic alternative based on the clustering of pseudopneumococcal pherotypes from different continents.

In conclusion, we propose that CSP sequence analysis can provide rapid accurate differentiation of *S. pseudopneumoniae* from closely related species *S. pneumoniae*, *S. mitis*, and *S. oralis*. With this in mind, we anticipate that some strains currently classified as atypical pneumococci can be identified as pseudopneumococci based on ComC sequencing. Use of ComC sequencing will simplify the gathering of data to understand the disease potential of this organism, which may be now emerging as a pathogen. To add to this observation, we would encourage other laboratories to sequence "atypical pneumococcal" strains to provide more sequences to confirm whether ComC may be used as a rapid marker for the identification of this emerging pathogen.

ACKNOWLEDGMENTS

We thank the Royal Free Microbiology Reference Laboratory staff for processing samples.

We thank the Royal Free Hospital Special Trustees for funding this project.

REFERENCES

1. **Abdeldaim GM, Stralin K, Olcen P, Blomberg J, Herrmann B.** 2008. Toward a quantitative DNA-based definition of pneumococcal pneumonia: a comparison of *Streptococcus pneumoniae* target genes, with special reference to the Spn9802 fragment. Diagn. Microbiol. Infect. Dis. **60**:143–150.
2. **Arbique JC, et al.** 2004. Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae* sp. nov. J. Clin. Microbiol. **42**:4686–4696.
3. **Balsalobre L, et al.** 2006. Molecular characterization of disease-associated streptococci of the mitis group that are optochin susceptible. J. Clin. Microbiol. **44**:4163–4171.
4. **Borek AP, Dressel DC, Hussong J, Peterson LR.** 1997. Evolving clinical problems with *Streptococcus pneumoniae*: increasing resistance to antimicrobial agents, and failure of traditional optochin identification in Chicago, Illinois, between 1993 and 1996. Diagn. Microbiol. Infect. Dis. **29**: 209–214.
5. **Carvalho Mda G, et al.** 2007. Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA. J. Clin. Microbiol. **45**:2460–2466.
6. **Dauga C.** 2002. Evolution of the *gyrB* gene and the molecular phylogeny of Enterobacteriaceae: a model molecule for molecular systematic studies. Int. J. Syst. Evol. Microbiol. **52**:531–547.
7. **El Aila NA, et al.** 2010. The development of a 16S rRNA gene based PCR for the identification of *Streptococcus pneumoniae* and comparison with four other species specific PCR assays. BMC Infect. Dis. **10**:104.
8. **Enright MC, Spratt BG.** 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. Microbiology **144**(Pt. 11):3049–3060.
9. **Fenoll A, Martinez-Suarez JV, Munoz R, Casal J, Garcia JL.** 1990. Identification of atypical strains of *Streptococcus pneumoniae* by a specific DNA probe. Eur. J. Clin. Microbiol. **9**:396–401.
10. **Fenoll A, Munoz R, Garcia E, de la Campa AG.** 1994. Molecular basis of the optochin-sensitive phenotype of pneumococcus: characterization of the genes encoding the F0 complex of the *Streptococcus pneumoniae* and *Streptococcus oralis* H(+)-ATPases. Mol. Microbiol. **12**:587–598.
11. **Hanage WP, et al.** 2005. Using multilocus sequence data to define the pneumococcus. J. Bacteriol. **187**:6223–6230.
12. **Harf-Monteil C, Granello C, Le Brun C, Monteil H, Riegel P.** 2006. Incidence and pathogenic effect of *Streptococcus pseudopneumoniae.* J. Clin. Microbiol. **44**:2240–2241.
13. **Havarstein LS, Coomaraswamy G, Morrison DA.** 1995. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in *Streptococcus pneumoniae.* Proc. Natl. Acad. Sci. U. S. A. **92**: 11140–11144.
14. **Havarstein LS, Gaustad P, Nes IF, Morrison DA.** 1996. Identification of the streptococcal competence-pheromone receptor. Mol. Microbiol. **21**: 863–869.
15. **Havarstein LS, Hakenbeck R, Gaustad P.** 1997. Natural competence in the genus *Streptococcus*: evidence that streptococci can change pherotype by interspecies recombinational exchanges. J. Bacteriol. **179**:6589–6594.
16. **Hoshino T, Fujiwara T, Kilian M.** 2005. Use of phylogenetic and phenotypic analyses to identify nonhemolytic streptococci isolated from bacteremic patients. J. Clin. Microbiol. **43**:6073–6085.
17. **Ing J, et al.** 11 January 2012. Characterization of non-typeable and atypical *Streptococcus pneumoniae* pediatric isolates from 1994–2010. J. Clin. Microbiol. doi:10.1128/JCM.05182-11.
18. **Johnston C, et al.** 2010. Detection of large numbers of pneumococcal virulence genes in streptococci of the mitis group. J. Clin. Microbiol. **48**: 2762–2769.
19. **Kacou-N'douba A, et al.** 2010. Emergence of optochin resistance among *S. pneumoniae* strains colonizing healthy children in Abidjan. Indian J. Med. Microbiol. **28**:80–81.
20. **Kaijalainen T, Rintamaki S, Herva E, Leinonen M.** 2002. Evaluation of gene-technological and conventional methods in the identification of *Streptococcus pneumoniae.* J. Microbiol. Methods **51**:111–118.
21. **Kawamura Y, Hou XG, Sultana F, Miura H, Ezaki T.** 1995. Determination of 16S rRNA sequences of *Streptococcus mitis* and *Streptococcus gordonii* and phylogenetic relationships among members of the genus *Streptococcus.* Int. J. Syst. Bacteriol. **45**:406–408.
22. **Kearns AM, et al.** 2000. Pneumolysin detection identifies atypical isolates of *Streptococcus pneumoniae.* J. Clin. Microbiol. **38**:1309–1310.
23. **Keith ER, Murdoch DR.** 2008. Antimicrobial susceptibility profile of *Streptococcus pseudopneumoniae* isolated from sputum. Antimicrob. Agents Chemother. **52**:2998.
24. **Keith ER, Podmore RG, Anderson TP, Murdoch DR.** 2006. Characteristics of *Streptococcus pseudopneumoniae* isolated from purulent sputum samples. J. Clin. Microbiol. **44**:923–927.
25. **Kilian M, et al.** 2008. Evolution of *Streptococcus pneumoniae* and its close commensal relatives. PLoS One **3**:e2683.
26. **Knoll B, Tleyjeh IM, Steckelberg JM, Wilson WR, Baddour LM.** 2007. Infective endocarditis due to penicillin-resistant viridans group streptococci. Clin. Infect. Dis. **44**:1585–1592.
27. **Kontiainen S, Sivonen A.** 1987. Optochin resistance in *Streptococcus pneumoniae* strains isolated from blood and middle ear fluid. Eur. J. Clin. Microbiol. **6**:422–424.
28. **Leegaard TM, et al.** 2010. Phenotypic and genomic characterization of pneumococcus-like streptococci isolated from HIV-seropositive patients. Microbiology **156**:838–848.
29. **Lejbkowicz F, Goldstein M, Hashman N, Cohn L.** 1999. Optochin-resistant *Streptococcus pneumoniae* isolated from a blood specimen. Clin. Microbiol. Newsl. **21**:72–73.
30. **Leung MH, Oriyo NM, Gillespie SH, Charalambous BM.** 2011. The

adaptive potential during nasopharyngeal colonisation of *Streptococcus pneumoniae*. Infect. Genet. Evol. **11**:1989–1995.
31. **Marron A, et al.** 2000. Serious complications of bacteremia caused by viridans streptococci in neutropenic patients with cancer. Clin. Infect. Dis. **31**:1126–1130.
32. **Martin-Galiano AJ, Balsalobre L, Fenoll A, de la Campa AG.** 2003. Genetic characterization of optochin-susceptible viridans group streptococci. Antimicrob. Agents Chemother. **47**:3187–3194.
33. **Muller-Graf CD, et al.** 1999. Population biology of *Streptococcus pneumoniae* isolated from oropharyngeal carriage and invasive disease. Microbiology **145**(Pt. 11):3283–3293.
34. **Munoz R, Fenoll A, Vicioso D, Casal J.** 1990. Optochin-resistant variants of *Streptococcus pneumoniae*. Diagn. Microbiol. Infect. Dis. **13**:63–66.
35. **Neeleman C, Klaassen CH, Klomberg DM, de Valk HA, Mouton JW.** 2004. Pneumolysin is a key factor in misidentification of macrolide-resistant *Streptococcus pneumoniae* and is a putative virulence factor of *S. mitis* and other streptococci. J. Clin. Microbiol. **42**:4355–4357.
36. **Nolan T, Hands RE, Ogunkolade W, Bustin SA.** 2006. SPUD: a quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. Anal. Biochem. **351**:308–310.
37. **Park HK, et al.** 2010. Identification of the *cpsA* gene as a specific marker for the discrimination of *Streptococcus pneumoniae* from viridans group streptococci. J. Med. Microbiol. **59**:1146–1152.
38. **Phillips G, Barker R, Brogan O.** 1988. Optochin-resistant *Streptococcus pneumoniae*. Lancet **ii**:281.
39. **Pikis A, Campos JM, Rodriguez WJ, Keith JM.** 2001. Optochin resistance in *Streptococcus pneumoniae*: mechanism, significance, and clinical implications. J. Infect. Dis. **184**:582–590.
40. **Shahinas D, et al.** 2011. Whole-genome sequence of *Streptococcus pseudopneumoniae* isolate IS7493. J. Bacteriol. **193**:6102–6103.
41. **Simoes AS, et al.** 2010. Highly penicillin-resistant multidrug-resistant pneumococcus-like strains colonizing children in Oeiras, Portugal: genomic characteristics and implications for surveillance. J. Clin. Microbiol. **48**:238–246.
42. **Simoes AS, Valente C, de Lencastre H, Sa-Leao R.** 2011. Rapid identification of noncapsulated *Streptococcus pneumoniae* in nasopharyngeal samples allowing detection of co-colonization and reevaluation of prevalence. Diagn. Microbiol. Infect. Dis. **71**:208–216.
43. **Sistek V, et al.** 25 October 2011. Development of a real-time PCR assay for the specific detection and identification of *Streptococcus pseudopneumoniae* using the *recA* gene. Clin. Microbiol. Infect. doi:10.1111/j.1469-0691.2011.03684.x.
44. **Suzuki N, et al.** 2005. Discrimination of *Streptococcus pneumoniae* from viridans group streptococci by genomic subtractive hybridization. J. Clin. Microbiol. **43**:4528–4534.
45. **Tamura K, et al.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**:2731–2739.
46. **Vestrheim DF, Gaustad P, Aaberge IS, Caugant DA.** 2011. Pherotypes of pneumococcal strains co-existing in healthy children. Infect. Genet. Evol. **11**:1703–1708.
47. **Walsh PS, Metzger DA, Higuchi R.** 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques **10**:506–513.
48. **Whatmore AM, Barcus VA, Dowson CG.** 1999. Genetic diversity of the streptococcal competence (com) gene locus. J. Bacteriol. **181**:3144–3154.
49. **Whatmore AM, et al.** 2000. Genetic relationships between clinical isolates of *Streptococcus pneumoniae*, *Streptococcus oralis*, and *Streptococcus mitis*: characterization of "atypical" pneumococci and organisms allied to *S. mitis* harboring *S. pneumoniae* virulence factor-encoding genes. Infect. Immun. **68**:1374–1382.

Pneumococcal sepsis and nasopharyngeal carriage

Bambos M. Charalambous and Marcus H. Leung

Purpose of review
Streptococcus pneumoniae (the pneumococcus) remains an important cause of invasive disease including bacteraemia. This review highlights recent findings related to pneumococcal bacteraemia, virulence factors, and multiple colonization, including strain competition, biofilm formation, and competence.

Recent findings
Countries with no vaccination programmes see vaccine serotypes still prevalent in disease, whereas the emergence of nonvaccine serotypes in nasopharyngeal carriage and invasive disease is seen in countries with conjugate vaccination in place. Co-colonizing strains are being uncovered with more sensitive methods, and may act synergistically or compete with each other for survival. Several factors such as iron uptake, quorum signalling and the *luxS* gene, involved in colonization and virulence, are discussed. The role of quorum sensing signalling molecules and formation of biofilms are being explored.

Summary
Epidemiological data suggest that the latest serotype-based conjugate vaccines should provide heightened protection, although serotype replacement is now being seen. Much remains to be elucidated about its biology during multiple colonization, when evolution and adaptation to its host take place. The modes of colonization (biofilm, intracellular or surface adherence to the mucosal epithelium), and whether organisms that cause invasive disease have attenuated ability to colonize the nasopharynx remain to be elucidated.

Keywords
invasive pneumococcal disease, nasopharyngeal colonization, pneumococcal vaccines, pneumococcus, serotypes, *Streptococcus pneumoniae*

INTRODUCTION

Streptococcus pneumoniae (the pneumococcus) is a common colonizer of the human nasopharynx. It also causes infections including otitis media and pneumonia, and invasive pneumococcal diseases (IPDs) such as meningitis and bacteraemia. The organism is estimated to be responsible for more than 1.6 million deaths per year, most of which occur in children and the elderly [1].

The organism is associated with more than 90 capsular serotypes differentiated by the composition and linkage of capsule polysaccharides [2–4]. These capsules are targets of pneumococcal vaccines, and the first pneumococcal conjugate vaccine (PCV7) is specific for seven of the most common serotypes (4, 6B, 9V, 14, 18C, 19F, and 23F). In 2010, a 13-valent vaccine was introduced, containing six additional serotypes (PCV7 serotypes and 1, 3, 5, 6A, 7F, and 19A). It is likely that the use of PCV13 will increase the protective coverage of the vaccine [5].

This review summarizes recent epidemiological findings pertaining to pneumococcal bacteraemia. In addition, studies investigating roles of virulence factors are described. Finally, we highlight recent studies on pneumococcal carriage with special attention to multiple colonization, and roles of quorum sensing systems and the formation of biofilms during asymptomatic colonization.

RECENT EPIDEMIOLOGICAL FINDINGS

Pneumococcal bacteraemia remains a concern in sub-Saharan Africa, where conjugate vaccination has only been implemented in a few countries. In Nigeria, where *S. pneumoniae* is not the leading causative agent in childhood bacteraemia, it is, however, the leading cause of death among bacteraemia cases [6]. In the absence of vaccines, PCV7 serotypes made up 75% of bacteraemia cases.

Centre for Clinical Microbiology, University College London Medical School, Royal Free Campus, Rowland Hill Street, London, UK

Correspondence to Dr Bambos M. Charalambous, Centre for Clinical Microbiology, UCL Medical School, Royal Free Campus, Rowland Hill Street, London NW3 2PF, UK. Tel: +44 207 794 0500; e-mail: b.charalambous@ucl.ac.uk

Curr Opin Pulm Med 2012, 18:222–227
DOI:10.1097/MCP.0b013e328352103b

KEY POINTS

- Given the high prevalence of bacteraemia cases caused by vaccine serotypes, the recent introduction of conjugate vaccines in Africa should provide a heightened level of protection, whereas other forms of treatment, such as HAART therapy, may also prevent pneumococcal diseases in immunocompromised groups.
- Sensitive methods in detecting multiple colonization in asymptomatic children indicate that a high proportion of hosts are colonized by heterogeneous strains together allowing the organisms to adapt rapidly by horizontal gene transfer.
- Interaction and competition between cells during colonization involve complex processes including bacteriocin production, biofilm formation, competence activation, and requirement for other surface proteins.

Similarly, serotypes 1, 6A, and 6B constitute over 40% of IPDs in children under 16 years old in Malawi [7]. In contrast, in areas where PCV7 has been introduced, a reduction in PCV7 serotype disease incidence was met with rises in IPD caused by nonvaccine serotypes (NVTs) including serotype 6C. A US study on paediatric isolates found high prevalence of 6C, with 44% of these isolated from blood [8▪]. Multilocus sequence typing (MLST) revealed that some of these serotype 6C sequence types were previously associated with vaccine serotypes 6B, 19A, and 23F, suggestive of capsular switching. A post-PCV7 study in Spain also identified NVTs 19A and 7F as the leading serotypes in bacteraemia [9].

Pneumococcal bacteraemia remains one of the most common complications in HIV-positive individuals. A recent study showed that pneumococcal bacteraemia prevalence decreased by over 90% within 4 years following highly active antiretroviral therapy (HAART) [10]. Remarkably, HAART was superior to conjugate vaccination in reducing bacteraemic HIV-positive patients after 2 years of treatment. It may be that for such high-risk groups, reduction of IPD incidence can be achieved by means other than pneumococcal vaccination.

Despite no clear sex differences in IPD incidence, animal studies revealed that disease susceptibility and mortality were greater in male mice compared with female mice in models of pneumonia and sepsis [11▪]. More male mice succumbed to infection by 48 h compared to female mice. In addition to significant weight loss and a reduction in body temperature, infected male mice were also associated with increased expression of cytokines such as interleukin (IL)-5, IL-7, regulated on activation normal T cell expressed and secreted, and KC (keratinocyte chemoattractant; keratinocyte-derived chemokine; formerly CXCL1). To our knowledge this is the first to compare sex with disease susceptibility.

It is known that some serotypes are more invasive than others. To assess and compare the disease potential of serotypes, a study was conducted in Israel to compare serotypes prevalent in the nasopharynx of healthy children with those admitted to hospital with community-acquired alveolar pneumonia (CAAP) [12]. The most common serotypes identified in the nasopharynx of CAAP children included 1, 5, 22F, 7F, 14, and 9V. Conversely, odds ratios (ORs) for colonization prevalence in CAAP compared to healthy individuals were less than 1 for serotypes 6A, 6B, 23A, and 35B, suggesting that these serotypes have lower disease potentials. Age-related differences in colonization were also examined, and serotypes 1 and 5 showed increased colonization with age, whereas serotypes 14 and 19F decreased, regardless of health status.

VIRULENCE/COLONIZATION FACTORS

An understanding of the biology of the organism may provide opportunities for new therapeutic platforms. Current epidemiological studies have been complemented by research into virulence factors that may also play roles in colonization.

The polysaccharide capsule enables the pneumococcus to evade opsonophagocytosis [13]. Given the important role of the spleen in host immunity, Lammers *et al.* [14] investigated the role of capsule in disease susceptibility in asplenic mice. All asplenic mice intranasally inoculated with high concentrations (4×10^7 CFU/ml) of acapsulate pneumococci survived after 48 h, compared to the death of over 85% of mice inoculated with capsulated D39 even at a two-logs lower concentration (4×10^5 CFU/ml).

Pneumolysin (encoded by the gene *ply*) is an important virulence factor that exerts its haemolytic activity by binding to the host cholesterol receptor and forming an oligomeric pore [15]. Recently, a variant *ply* allele (*ply*4496) in a hypervirulent lineage of serotype 1 showed reduced haemolytic activity, and while strain D39 expressing this allele was attenuated in murine sepsis, higher bacterial count in blood was detected compared to the wild type, suggesting that reduced haemolytic activity confers a growth advantage in blood [16].

Immune evasion is a principal mechanism for pneumococcal survival. The pneumococcal surface protein C (PspC) binds to the host complement factor H (fH), escaping recognition of the alternative complement pathway. The potential of the

fH-binding region of PspC as a vaccine target was investigated recently [17]. Immunization with a booster in mice increased IgG levels of this epitope, and intravenous injection of the homologous strain into vaccinated mice showed increased survival compared to nonimmunized mice, accompanied with increased C3 binding, lowered fH binding, and increased susceptibility to phagocytosis. However, challenge with heterogeneous strains showed lowered levels of protection, possibly dependent on the variation of PspC amino acid sequences between the vaccine antigen and that of the challenge strain. Given the high sequence diversity of *pspC* [18], it is unlikely that PspC alone as a vaccine antigen can provide protection from across different pneumococcal strains.

Metal ions are known to be required for pneumococcal cellular function and pathogenicity [19,20]. Bayle *et al.* [21] reported the requirement of zinc in normal pneumococcal cellular function, as mutations of two components involved in zinc transport, AdcA and AdcAII, were associated with abnormalities in cell division, extended lag phase, and attenuation in murine colonization, pneumonia, and sepsis.

NASOPHARYNGEAL COLONIZATION

Infections caused by the pneumococcus are rare compared to colonization. During colonization, microevolution and adaptation arise by horizontal gene transfer (HGT) of capsular genes and genes encoding antibiotic resistance determinants and surface antigens under selection [22,23]. This results in a supragenome (or pan-genome) that is larger than the genome of any single organism [24–26]. Therefore, the co-colonization (or multiple colonization) of pneumococcal strains increases the repertoire of genes available for adaptation and virulence, enhancing its survival as a species [27].

Multiple colonization was documented as early as the 1930s [28], but few studies since have determined multiple colonization by serotyping multiple colonies on blood agar, and even fewer studies accompanied serotype results with genetic analysis of strains, or performed serotyping with more sensitive DNA-based methods. This results in an underestimation of strain diversity. Recently, two studies [29▪▪,30▪▪] addressed this issue using DNA-based serotyping methods in association with MLST to investigate strain diversity in the nasopharynx of Gambian and Tanzanian children (Table 1). Cohort differences and detection methods used may account for some of the difference seen in co-colonization rates. Leung *et al.* [30▪▪] observed the presence of unrelated strains of the same serotype, highlighting the underestimation of co-colonizations by characterizing serotype alone. The same study also documented the simultaneous colonization of strains with different susceptibility to penicillin and trimethoprim–sulfamethoxazole. Both Gambian and Tanzanian studies found a large proportion of sequence types exclusive to Africa, as well as novel sequence types. They hypothesize that both the community prevalence of multiple colonization and the extent of strain diversity within the nasopharynx encourage HGT between strains, leading to frequent assortments of MLST gene fragments giving rise to these novel sequence types. Donkor *et al.* [29▪▪] also reported evidence of capsular switching.

Detection of co-colonizations requires sensitive methods. Three methods were compared in Thai children (Table 1) [31▪▪]. In addition to high rates of co-colonization, microarray detected a high prevalence of nontypable pneumococci. Nontypable pneumococci may comprise acapsulate strains, which had traditionally thought to be rare colonizers due to the loss of protective capsule. However, the use of DNA-based methods may increase the likelihood of detecting nontypables by as much as three-fold [32▪▪]. Genetic analysis indicated that most nontypables were associated with ST304, a sequence type associated with conjunctivitis, blood infections, and multidrug resistance [32▪▪,33]. These data should be treated cautiously as DNA-based methods do not reveal only current co-colonizations.

Table 1. Recent studies of multiple colonization using sensitive detection methods

Country	Age group (months)	Method[a]	Prevalence (%)	No. max serotype	Reference[b]
The Gambia	≤24	Microarray and CS	19	3	[29▪▪]
Tanzania	≤72	CS of 20 colonies and sequence-based sequetyping	12	5	[30▪▪]
Thailand	≤24	CS on three colonies	11	2	[31▪▪]
		Sweep colony with latex agglutination	43	4	[31▪▪]
		Microarray	48	9	[31▪▪]

[a]CS, conventional serotyping by Quellung reaction.
[b]In the study by Turner *et al.* [31▪▪], three methods were performed on the same cohort.

The method of Simoes *et al.* [32▪▪] also distinguished pneumococci from *S. pseudopneumoniae*. Identification of *S. pseudopneumoniae* is difficult due to its strong genotypic and phenotypic resemblance to nontypable pneumococci. However, *S. pseudopneumoniae* appears to be differentiable from nontypable pneumococci by the absence of *lytA* targeted by primers described. It must be noted only two pseudopneumococcal strains were included in this study, and *lytA*-positive strains of *S. pseudopneumoniae* have been reported [34].

The pneumococcus shares the same ecological niche with *Haemophilus influenzae*, *Moraxella catarrhalis*, and *Staphylococcus aureus*. The prevalence of multiple colonization of these organisms in Gambian infants was assessed using PCR targeting species-specific regions [35]. More than 90% of children were detected to be colonized by at least one of these four opportunistic pathogens by one year of age, and may acquire these species as early as 3 weeks old. There also appears to be a strong and more modest correlation between pneumococcal colonization and that of *H. influenzae* and *M. catarrhalis*, respectively, in which colonization point-prevalence for all three organisms from newborn to 27 weeks increased. Remarkably, the opposite was observed for *S. aureus*, in which case colonization prevalence declined by approximately 80% in the same period. Such information is crucial in understanding the effects of pneumococcal vaccination on colonization of other potential pathogens.

BACTERIOCINS, BIOFILM FORMATION, AND COMPETENCE

Competition between pneumococcal strains is mediated by small antimicrobial peptides (bacteriocins). Bacteriocin secretion is modulated by the two-component system encoded by the *blp* locus consisting of an ABC transporter (*blpAB*), a peptide pheromone (*blpC*), a histidine kinase receptor (*blpH*), and response regulator (*blpR*). Pneumococci are protected by their own bacteriocins by the synthesis of immunity proteins. A lineage of high number of clinical isolates was detected to contain a 4-bp repeat insertion sequence in *blpA* that renders these strains unable to secrete BlpC, but they are still responsive to BlpC secreted by other strains, and could produce bacteriocin and immunity proteins [36]. These 'cheater' strains are thought to have acquired the insertion via HGT, and are likely to confer a survival advantage by benefiting from bacteriocin production from nearby strains. Son *et al.* [36] hypothesizes that such an advantage drives the diversification of BlpC so that cheater strains do not respond to heterologous pheromones.

Pneumococci are thought to colonize as a biofilm, a complex polymicrobial community initiated by adherence to the mucosal epithelium. Biofilm formation is related to competence, a transient physiological state in which pneumococci can acquire DNA. Similar to *blp* activation, competence in pneumococci is modulated by the two-component system encoded by the *com* locus, and the signalling molecule involved in competence activation (ComC) exists in allelic variants (or pherotypes) [37]. Competence is also responsible for the activation of fratricide, when a pherotype lyses noncompetent strains. The co-colonization of different pherotypes was documented for the first time in asymptomatic children [38▪▪]. The authors rationalized that the co-existence of different pherotypes provides evidence that fratricide appears to have little effect on competing pherotypes.

Competence leads to the expression of stress response genes [39], and was shown to be activated in response to erroneous protein translation and antibiotics [40]. Pneumococcal HtrA (heat-shock-induced, surface-associated serine protease) appears to modulate this regulation. Whereas the mechanism of how HtrA influences activation of competence is unclear, CSP (competence stimulating peptide) may be degraded by HtrA. It is unknown whether host responses that affect pneumococcal protein misfolding also affect competence [40].

The role of competence in biofilm formation depends on the biofilm model [41▪]. In microtitre plates, pherotype-specific CSP seemed to affect the late phase of biofilm formation for both exponential and stationary-phase cells, as *comC* and *comD* deletion mutants had significantly reduced biofilm cell count by 24 h. Interestingly, the addition of a different exogenous pherotype did not support the maintenance of a biofilm, despite previous indications of possible pherotype cross-talk [42]. CSP does not seem to have a role in continuous biofilm formation, whereas the lack of capsule was associated with increased biofilm thickness, cell count, and surface area [41▪,43].

Although adhesion and biofilm formation are thought to occur in colonization, it is uncertain whether this occurs with invasive strains. Comparison of different serotypes captured from IPD revealed that biofilm development is a general pneumococcal property [43]. Some serotypes (serotypes 6A, 6B, 7F) formed better biofilms than others (serotype 3). However, the genetic background and antibiotic susceptibility may also contribute to biofilm formation.

Competence, quorum sensing, and biofilm formation have been linked to the *luxS* gene, in which early biofilm formation was associated with maximal

expression (up to 300-fold) of *luxS* during mid-log phase, with increases in *ply* and *lytA* mRNA levels [44]. Strains with inactivated *luxS* had reduced biofilm formation, with a lower count of attached cells and a lack of biofilm architecture. The involvement of LuxS in biofilm formation was further supported by Trappetti *et al.* [45▪▪], when LuxS activity and biofilm growth appeared to be dependent on acquisition of extracellular iron. Mutants deficient in *luxS* may have weakened ability to acquire iron, as a reduced expression of the iron acquisition lipoprotein PiuA was seen. The authors further associated *luxS* expression and availability of iron to up-regulation of genes involved in competence and fratricide. These current studies on biofilm development and its regulation indicate a complex process involving multiple cellular signalling pathways and the expression of a wide array of proteins in response to different environmental factors.

CONCLUSION

Pneumococcal bacteraemia is a common cause of death in African children. With the recent introduction of PCV13, a reduction in IPD caused by vaccine serotypes is anticipated. However, replacement serotypes may follow, as seen in countries with PCV7 implemented. Increasing our understanding of virulence/colonization factors may be beneficial in unravelling pneumococcal pathogenicity. Multiple pneumococcal colonization is an emerging concept, as shown by efforts to increase the sensitivity of detection. Finally, understanding the interaction of these different strains during colonization, such as the biofilm mode of growth, competition and communication between species, may provide new therapeutic platforms to prevent pneumococcal disease in the face of the global increase in antibiotic resistance and serotype replacement post vaccination.

Acknowledgements

Marcus Leung has been partly supported by the Royal Free Charity.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ▪ of special interest
- ▪▪ of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (pp. 283–284).

1. World Health Organization. Pneumococcal conjugate vaccine for childhood immunization: WHO position paper. Wkly Epidemiol Rec 2007; 82:93–104.
2. Bentley SD, Aanensen DM, Mavroidi A, *et al.* Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. PLoS Genet 2006; 2:e31.
3. Calix JJ, Nahm MH. A new pneumococcal serotype, 11E, has a variably inactivated wcjE gene. J Infect Dis 2010; 202:29–38.
4. Park IH, Park S, Hollingshead SK, Nahm MH. Genetic basis for the new pneumococcal serotype, 6C. Infect Immun 2007; 75:4482–4489.
5. Gladstone RA, Jefferies JM, Faust SN, Clarke SC. Continued control of pneumococcal disease in the UK: the impact of vaccination. J Med Microbiol 2011; 60:1–8.
6. Obaro S, Lawson L, Essen U, *et al.* Community acquired bacteremia in young children from central Nigeria: a pilot study. BMC Infect Dis 2011; 11:137.
7. Cornick JE, Everett DB, Broughton C, *et al.* Invasive Streptococcus pneumoniae in children, Malawi, 2004–2006. Emerg Infect Dis 2011; 17:1107–1109.
8. ▪ Green MC, Mason EO, Kaplan SL, *et al.* Increase in prevalence of *Streptococcus pneumoniae* serotype 6C at eight children's hospitals in the United States from 1993 to 2009. J Clin Microbiol 2011; 49:2097–2101.

This work encompasses prevaccine and postvaccine periods in illustrating the phenomenon of serotype replacement with attention to a recently characterized serotype.

9. Picazo J, Ruiz-Contreras J, Casado-Flores J, *et al.* Relationship between serotypes, age, and clinical presentation of invasive pneumococcal disease in Madrid, Spain, after introduction of the 7-valent pneumococcal conjugate vaccine into the vaccination calendar. Clin Vaccine Immunol 2011; 18:89–94.
10. Muyanja SZ, Larke N, Rutebarika D, *et al.* Decreasing trends of bacteraemia among HIV-infected Ugandan adults: incidence, aetiology, clinical outcomes and effect of antiretroviral therapy in a semi-urban setting (2000–2008). Trop Med Int Health 2011; 16:756–765.
11. ▪ Kadioglu A, Cuppone AM, Trappetti C, *et al.* Sex-based differences in susceptibility to respiratory and systemic pneumococcal disease in mice. J Infect Dis 2011; 204:1971–1979.

First publication that has indicated a sex difference in disease susceptibility.

12. Greenberg D, Givon-Lavi N, Newman N, *et al.* Nasopharyngeal carriage of individual Streptococcus pneumoniae serotypes during pediatric pneumonia as a means to estimate serotype disease potential. Pediatr Infect Dis J 2011; 30:227–233.
13. Hyams C, Camberlein E, Cohen JM, *et al.* The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. Infect Immun 2010; 78:704–715.
14. Lammers AJ, de Porto AP, Florquin S, *et al.* Enhanced vulnerability for *Streptococcus pneumoniae* sepsis during asplenia is determined by the bacterial capsule. Immunobiology 2011; 216:863–870.
15. Rossjohn J, Gilbert RJ, Crane D, *et al.* The molecular mechanism of pneumolysin, a virulence factor from *Streptococcus pneumoniae*. J Mol Biol 1998; 284:449–461.
16. Harvey RM, Ogunniyi AD, Chen AY, Paton JC. Pneumolysin with low hemolytic activity confers an early growth advantage to *Streptococcus pneumoniae* in the blood. Infect Immun 2011; 79:4122–4130.
17. Ricci S, Janulczyk R, Gerlini A, *et al.* The factor H-binding fragment of PspC as a vaccine antigen for the induction of protective humoral immunity against experimental pneumococcal sepsis. Vaccine 2011; 29:8241–8249.
18. Iannelli F, Oggioni MR, Pozzi G. Allelic variation in the highly polymorphic locus pspC of *Streptococcus pneumoniae*. Gene 2002; 284:63–71.
19. Brown JS, Gilliland SM, Ruiz-Albert J, Holden DW. Characterization of pit, a *Streptococcus pneumoniae* iron uptake ABC transporter. Infect Immun 2002; 70:4389–4398.
20. Morona JK, Morona R, Miller DC, Paton JC. *Streptococcus pneumoniae* capsule biosynthesis protein CpsB is a novel manganese-dependent phosphotyrosine-protein phosphatase. J Bacteriol 2002; 184:577–583.
21. Bayle L, Chimalapati S, Schoehn G, *et al.* Zinc uptake by *Streptococcus pneumoniae* depends on both AdcA and AdcAII and is essential for normal bacterial morphology and virulence. Mol Microbiol 2011; 82:904–916.
22. Croucher NJ, Harris SR, Fraser C, *et al.* Rapid pneumococcal evolution in response to clinical interventions. Science 2011; 331:430–434.
23. Hiller NL, Ahmed A, Powell E, *et al.* Generation of genic diversity among Streptococcus pneumoniae strains via horizontal gene transfer during a chronic polyclonal pediatric infection. PLoS Pathog 2010; 6.
24. Donati C, Hiller NL, Tettelin H, *et al.* Structure and dynamics of the pangenome of *Streptococcus pneumoniae* and closely related species. Genome Biol 2010; 11:R107.
25. Hiller NL, Janto B, Hogg JS, *et al.* Comparative genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the pneumococcal supragenome. J Bacteriol 2007; 189:8186–8195.
26. Shen K, Gladitz J, Antalis P, *et al.* Characterization, distribution, and expression of novel genes among eight clinical isolates of *Streptococcus pneumoniae*. Infect Immun 2006; 74:321–330.
27. Harvey RM, Stroeher UH, Ogunniyi AD, *et al.* A variable region within the genome of *Streptococcus pneumoniae* contributes to strain-strain variation in virulence. PLoS One 2011; 6:e19650.
28. Gundel M, Okura G. Untersuchungen über das gleichzeitige Vorkommen mehrerer Pneumokokkentypen bei Gesunden und ihre Bedeutung für die Epidemiologie. Med Microbiol Immunol 1933; 114:678–704.

29. Donkor ES, Bishop CJ, Gould KA, *et al.* High levels of recombination among ■■ Streptococcus pneumoniae isolates from the Gambia. MBio 2011; 2:e00040-11.
This publication revealed a high number of novel sequence types, suggested to arise from gene exchange which is enhanced in areas of high rates of co-colonization.
30. Leung MH, Oriyo NM, Gillespie SH, Charalambous BM. The adaptive ■■ potential during nasopharyngeal colonisation of *Streptococcus pneumoniae*. Infect Genet Evol 2011; 11:1989–1995.
This study uncovered a high extent of genetic diversity within a colonization, with co-colonization of antibiotic-susceptible and nonsusceptible strains, and unrelated strains of the same serotype, indicating that serotype alone underestimates true heterogeneity of carriage.
31. Turner P, Hinds J, Turner C, *et al.* Improved detection of nasopharyngeal ■■ cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. J Clin Microbiol 2011; 49:1784–1789.
Using microarray, the authors have uncovered up to nine serotypes colonizing the same host at one time, with fold-changes of increase compared to traditional culture methods.
32. Simoes AS, Valente C, de Lencastre H, Sa-Leao R. Rapid identification ■■ of noncapsulated *Streptococcus pneumoniae* in nasopharyngeal samples allowing detection of co-colonization and reevaluation of prevalence. Diagn Microbiol Infect Dis 2011; 71:208–216.
This publication emphasized the finding that colonization by nontypable pneumococci is higher than previously thought, and may play an important role in pneumococcal adaptation.
33. Berron S, Fenoll A, Ortega M, *et al.* Analysis of the genetic structure of nontypeable pneumococcal strains isolated from conjunctiva. J Clin Microbiol 2005; 43:1694–1698.
34. Carvalho Mda G, Tondella ML, McCaustland K, *et al.* Evaluation and improvement of real-time PCR assays targeting *lytA*, *ply*, and *psaA* genes for detection of pneumococcal DNA. J Clin Microbiol 2007; 45:2460–2466.
35. Kwambana BA, Barer MR, Bottomley C, *et al.* Early acquisition and high nasopharyngeal co-colonisation by *Streptococcus pneumoniae* and three respiratory pathogens amongst Gambian new-borns and infants. BMC Infect Dis 2011; 11:175.
36. Son MR, Shchepetov M, Adrian PV, *et al.* Conserved mutations in the pneumococcal bacteriocin transporter gene, *blpA*, result in a complex population consisting of producers and cheaters. MBio 2011; 2:e00179-11.
37. Whatmore AM, Barcus VA, Dowson CG. Genetic diversity of the streptococcal competence (*com*) gene locus. J Bacteriol 1999; 181:3144–3154.
38. Vestrheim DF, Gaustad P, Aaberge IS, Caugant DA. Pherotypes of pneu- ■■ mococcal strains co-existing in healthy children. Infect Genet Evol 2011; 11:1703–1708.
This is the first publication highlighting the colonization of multiple competence pherotypes.
39. Peterson SN, Sung CK, Cline R, *et al.* Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. Mol Microbiol 2004; 51:1051–1070.
40. Stevens KE, Chang D, Zwack EE, Sebert ME. Competence in *Streptococcus pneumoniae* is regulated by the rate of ribosomal decoding errors. MBio 2011; 2:e00071-11.
41. Trappetti C, Gualdi L, Di Meola L, *et al.* The impact of the competence quorum ■ sensing system on *Streptococcus pneumoniae* biofilms varies depending on the experimental model. BMC Microbiol 2011; 11:75.
Findings in this publication allows a more catagorical assessment of previous works on roles of competence on biofilm formation based on the different biofilm models employed.
42. Iannelli F, Oggioni MR, Pozzi G. Sensor domain of histidine kinase ComD confers competence pherotype specificity in *Streptococcus pneumoniae*. FEMS Microbiol Lett 2005; 252:321–326.
43. Camilli R, Pantosti A, Baldassarri L. Contribution of serotype and genetic background to biofilm formation by *Streptococcus pneumoniae*. Eur J Clin Microbiol Infect Dis 2011; 30:97–102.
44. Vidal JE, Ludewick HP, Kunkel RM, *et al.* The LuxS-dependent quorum-sensing system regulates early biofilm formation by *Streptococcus pneumoniae* strain D39. Infect Immun 2011; 79:4050–4060.
45. Trappetti C, Potter AJ, Paton AW, *et al.* LuxS mediates iron-dependent biofilm ■■ formation, competence, and fratricide in *Streptococcus pneumoniae*. Infect Immun 2011; 79:4550–4558.
The finding that LuxS plays roles in affecting biofilm formation, competence, and fratricide may soon allow a greater understanding of the interconnection of these complex processes.